Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

The Role of Neural Replay in Shaping and Utilizing Neural Representations Supporting Learning, Memory, and Decision-Making

Dissertation zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)

an der Universität Hamburg,
Fakultät für Psychologie und Bewegungswissenschaft,
Institut für Psychologie

vorgelegt von

Fabian M. Renz

Hamburg, 2026

Tag der Disputation: 1.9.2026

Promotionsprüfungsausschuss:

Vorsitzende/r:	Prof. Dr. Jan Wacker
1. Dissertationsgutachter:	Prof. Dr. Nicolas Schuck
2. Dissertationsgutachter:	Prof. Dr. Christian Doeller
1. Disputationsgutachter/in:	Prof. Dr. Anja Riesel
2. Disputationsgutachter/in:	Prof. Daphna Shohamy, PhD

Abstract

Flexible generalization depends on internal representations that capture the latent structure of the environment and can be dynamically updated as that structure changes. This thesis investigates the role of neural replay, the offline reactivation of sequential neural patterns, in building, refining, and leveraging such representations across wakefulness and sleep.

The first empirical project combined functional magnetic resonance imaging (fMRI) with computational modelling to study how replay supports structure learning and generalization. Fifty-two participants performed a sequential decision task in which multiple options shared hidden reward sources. A latent cause inference model captured participants' discovery of this structure and outperformed standard reinforcement learning models. Backward replay in visual cortex selectively targeted unobserved sequences sharing the inferred latent reward structure, increased as participants discovered that structure, and tracked trial-by-trial non-local value updating from the computational model. When the latent structure changed on a second day, replay reorganised accordingly, including a directional shift for newly generalizable paths.

The second empirical project examined how sleep transforms the neural basis of sequential memory retrieval. Participants learned hierarchically organized sequences and were tested before and after overnight sleep. Sequential reactivation during retrieval shifted in regional profile across consolidation: visual cortex dominated before sleep, whereas parietal cortex emerged as the primary locus after sleep, with replay strength predicting retrieval accuracy. This reorganisation is consistent with the systems consolidation hypothesis that sleep drives the redistribution of memory representations toward cortical networks.

A methodological book chapter situated these findings within the broader landscape of tools for measuring neural plasticity in humans. Together, the three projects converge on a view of replay as an active computational mechanism sampling from an internal generative model supporting multiple functions including rapid value propagation over inferred structure during wakefulness, and driving cortical integration of memory representations during sleep.

Contents

Part I

Overview

> *"Behaviour must be flexible as a function of the environment. If a system cannot make itself respond in whatever way is needed, it hardly can be mind-like. [...] We must be able to learn from the environment, not occasionally but continuously, and not about a few specifics but everything and every way under the sun."*
> — Allen Newell, *Unified Theories of Cognition* (1990), p. 19

Introduction

Adaptive behavior is a defining hallmark of intelligent agents. To navigate a complex, ever-changing world, agents must be capable of flexible adaptation, continuously learning from experience and utilizing that knowledge to guide future behavior [Botvinick and Toussaint, 2012]. Yet accumulating experience is not enough. As Roger Shepard famously argued, "the first law of psychology should be one of generalization": the capacity to extend past experience to entirely novel situations is essential for coping with circumstances never before encountered [Shepard, 1987]. Intelligent agents therefore depend not only on the accumulation of knowledge, but on the ability to generalize it in principled ways, enabling adaptive behavior even when the present differs from the past.

Generalization of this kind requires more than storing additional information. To act appropriately in new situations, an agent must infer latent structure, that is, the hidden regularities, causal relationships, and abstract principles underlying observed events, and use this structure to abstract away from situation-specific details toward aspects that transfer across contexts [Radulescu et al., 2021, Gershman and Niv, 2010, Niv, 2019]. As evidence accumulates, internal models of the world are reshaped so that previously separate experiences can be reinterpreted as instances of a shared underlying structure [Tse et al., 2007, Gershman et al., 2010]. It is this capacity to revise, reorganize, and enrich internal models that allows past experience to inform future behavior in flexible and systematic ways.

This thesis asks how neural replay, the offline reactivation of sequential neural activity patterns [Wittkuhn et al., 2021], contributes to the learning and leveraging of such structured representations. Each project approaches this question from a different angle, together spanning the arc from the online leveraging of learned structure during wakefulness to offline representational refinement during sleep, and from computational theory to neural measurement.

This thesis is organized into four parts, each structured into chapters. Part I, the present part, contains an introductory chapter and the current overview, which foreshadows the main points of the subsequent thesis. The remaining three parts develop the theoretical foundations, present the empirical and methodological contributions, and offer a synthesis.

Part II: Theoretical Foundations

Part II provides the theoretical and methodological background for the three contributions of this thesis. Each introductory chapter is paired with a specific contribution in Part III, develop-

ing the concepts and prior work needed to motivate the empirical questions addressed there.

Chapters 1 and 2 together introduce the first empirical project, which investigates how replay during wakefulness supports the learning of latent task structure and its use for generalization, value updating, and flexible decision-making. Chapter 1 develops the computational dimension: it reviews reinforcement learning as a framework for value-based decision-making, examines the problem of learning internal task representations, introduces cognitive maps as structured relational models, and develops latent cause inference as a formal account of how agents discover hidden task structure and use it for non-local value propagation. Chapter 2 turns to the neural dimension: it defines replay and distinguishes it from broader notions of reactivation, reviews the major empirical findings on replay in animals and humans, surveys methods for detecting replay in human neuroimaging data, and connects these findings to the computational theories introduced in Chapter 1 to motivate the specific hypotheses tested in the accompanying empirical paper.

Chapter 3 introduces the second empirical project, which shifts from online structure learning to offline representational refinement. It reviews the systems consolidation hypothesis, the evidence for replay as its mechanistic implementation during sleep, and the emerging understanding of how consolidation transforms the content and anatomical locus of memory representations.

Chapter 4 provides a broader methodological perspective by surveying modern tools for investigating neural plasticity in humans. It covers structural and functional neuroimaging techniques, noninvasive brain stimulation, and the study designs needed to draw valid inferences about experience-dependent change. This chapter situates the methods deployed across the thesis within the wider landscape of tools available for measuring neural change in the human brain.

Part III: Empirical Contributions

Part III presents the three contributions of the thesis.

The first contribution is an empirical paper combining fMRI with computational modelling to study how replay supports value generalization during wakefulness. Fifty-two participants performed a sequential decision task in which multiple options shared hidden reward sources. A latent cause inference model captured participants' discovery of this structure and outperformed standard reinforcement learning models. Backward replay in visual cortex selectively targeted unobserved sequences sharing the inferred latent reward structure, increased as participants discovered that structure, and tracked trial-by-trial non-local value updating from the computational model. When the latent structure changed on a second day, replay reorganised accordingly.

The second contribution examines how sleep transforms the neural basis of sequential memory retrieval. Participants learned hierarchically organised sequences and were tested before

and after overnight sleep. Sequential reactivation during retrieval shifted in regional profile across consolidation: visual cortex dominated before sleep, whereas parietal cortex emerged as the primary locus after sleep, with replay strength predicting retrieval accuracy. This reorganisation is consistent with the systems consolidation hypothesis.

The third contribution is a methodological book chapter that surveys neuroscience methods for investigating brain plasticity, published in *The Oxford Handbook of Cognitive Enhancement and Brain Plasticity*. It reviews structural and functional neuroimaging techniques and discusses the design considerations needed to measure experience-dependent neural change, providing the methodological context for the preceding empirical work.

Part IV: Synthesis

The General Discussion integrates the findings across projects, organized around two threads. The first concerns function: what the results reveal about the computational roles of replay across brain states, and how they relate to competing accounts of replay as serving value updating, planning, consolidation, or structure learning. The second concerns method: the assumptions, strengths, and failure modes of fMRI-based replay detection, drawing on simulation work that characterizes the properties of the sequential slope ordering analysis used throughout the thesis. A central finding that emerges across the two empirical projects is that replay's function shifts with brain state: during wakefulness it supports rapid value propagation across inferred latent structure, while across sleep it is associated with the cortical redistribution of sequential memory representations. The common thread is replay as a model-based sampling process through which the brain builds, refines, and leverages structure for flexible, adaptive behavior.

Part II

Theoretical Foundations

Chapter 1

Representation Learning and Latent Structure in Value-Based Decision-Making

The overview established that flexible, generalizable behavior depends on the brain's capacity to learn and reorganize internal representations, and that neural replay is hypothesized to take a central role in the process of forming and leveraging such representations. This chapter develops the computational foundations needed to understand that claim. It examines how agents discover the structure of their environment, how that structure is represented and used to support generalization and value-based decision-making, and how offline mechanisms such as replay can propagate information across an inferred representational space. The neural instantiation of these processes is then taken up in the following chapter.

1.1 Reinforcement Learning as a Framework for Learning-Related Plasticity

Humans learn from a wide variety of experiences: from rewards and punishments, but also from observation, instruction, statistical regularities, episodic memories, and explicit reasoning [Anderson, 2000]. No single framework fully captures this diversity of learning processes. Nevertheless, reinforcement learning (RL) has emerged as a particularly influential framework for understanding how behaviour is shaped by the consequences of actions, and it accounts for an important subset of learning phenomena in both humans and other animals [Sutton and Barto, 2018].

In its classical formulation, RL describes how behaviour can be adapted through trial and error by tracking the outcomes of actions in terms of rewards and punishments. The key idea is that agents learn to assign *values* to states and actions, reflecting their expected future reward, and adjust behaviour so as to maximise this expected return over time. RL models thus provide a normative account of how feedback can drive gradual improvement in performance. This normative status rests on formal optimality guarantees: the Bellman optimality equation defines

the value of each state-action pair recursively, as the immediate reward plus the discounted value of the best action available in the resulting state, thereby characterising the optimal value function as a fixed point of this recursion. Algorithms such as Q-learning exploit this structure by iteratively updating value estimates from sampled experience, and are proven to converge to this optimal value function under standard assumptions, establishing a principled benchmark against which actual behaviour can be evaluated [Watkins and Dayan, 1992, Sutton and Barto, 2018]. The idea that learning is driven by discrepancies between expected and actual outcomes has deep roots. In their influential model of classical conditioning, Rescorla [1972] proposed that an animal learns to associate a cue with an outcome only to the extent that the outcome is surprising: if the outcome is already well predicted, little further learning occurs. This simple principle, that learning is proportional to prediction error, provided one of the earliest formalisations of error-driven learning and remains foundational to modern RL. In RL, the concept is generalised as the *reward prediction error*: the difference between received and expected outcomes, signalling whether events turned out better or worse than anticipated. These prediction errors drive incremental updates of value estimates, ensuring that future expectations better match experienced outcomes. Temporal-difference learning offers a particularly influential class of such algorithms, in which prediction errors are computed at each time step and used to update values based on local information about successive states and rewards [Sutton and Barto, 2018].

A major reason why RL has become so prominent in neuroscience is that these computational notions map onto identifiable neural signals. A landmark demonstration comes from recordings of midbrain dopamine neurons in monkeys performing a conditioning task [Schultz et al., 1997]. Early in learning, when a monkey receives an unexpected juice reward, dopamine neurons fire a sharp burst at the moment of reward delivery. As the animal learns that a visual cue reliably predicts the reward, this burst shifts: dopamine neurons now fire at the onset of the predictive cue rather than at reward delivery, consistent with the prediction error having moved to the earliest unpredicted event. Critically, if the expected reward is then omitted, dopamine activity dips below baseline at the time the reward would have occurred, a negative prediction error. This pattern, excitation for positive surprises, suppression for negative ones, and silence when expectations are met, closely matches the formal properties of a temporal-difference reward prediction error [Schultz, 1998, Montague et al., 1996]. Relatedly, activity in striatal and prefrontal regions often scales with model-derived prediction errors and value signals during human and animal learning tasks [e.g., O'Doherty et al., 2004, Daw et al., 2011, Hare et al., 2009]. These findings have motivated the view that RL-like mechanisms, implemented via dopamine-modulated synaptic plasticity in corticostriatal circuits, contribute substantially to learning from reward and punishment.

In this way, RL offers a powerful language for linking behaviour and neural signals: changes in synaptic strength are driven by error signals, which update value estimates and in turn shape future choices. Importantly, however, RL is best understood as a broad framework for the prob-

lem of learning from interaction. The framework defines the problem, a set of core concepts (states, actions, rewards, and values), and a family of algorithms for solving it, but it does not claim to encompass all of cognition. Phenomena such as learning from instruction, attentional selection, or relational reasoning do not contradict RL, but complement and extend it [Dayan and Niv, 2008, Gershman and Daw, 2017]. Some of these, including attention and memory, have since been incorporated into more recent RL architectures [Mnih et al., 2014, Graves et al., 2016], illustrating the framework's extensibility. RL is therefore best viewed as one component in a broader mosaic of learning systems, particularly well-suited to capturing incremental, feedback-driven adaptation.

Moreover, in its classical, "model-free" form, RL primarily specifies *how* values are updated given a set of states and actions, rather than *how* those states and their relationships are represented in the first place. The structure of the environment is typically assumed to be given, and prediction errors are used to adjust values within this fixed representational space [Niv, 2019]. This is a consequential limitation. As the remainder of this chapter argues, the choice of state representation fundamentally shapes what can be learned and how flexibly behaviour can adapt when contingencies change. The following sections therefore shift the focus from value updating alone to the problem of discovering and representing the latent structure of the environment.

1.2 Formalising Value Learning in Markov Decision Processes

In its standard formulation, reinforcement learning formalises decision-making as a Markov decision process (MDP) comprising a set of states S, actions A, a transition function $P(s' \mid s, a)$ specifying how actions change the state of the environment, a reward function $r(s, a)$, and a discount factor γ that determines the relative weighting of immediate and future rewards [Sutton and Barto, 2018]. A *policy* $\pi(a \mid s)$ defines how likely an agent is to choose each action in a given state. The central quantity of interest is the *value function*, which assigns to each state (or state–action pair) its expected long-term return under a particular policy. For example, the state-value function $V^\pi(s)$ is defined as the expected discounted sum of future rewards when starting in state s and following policy π thereafter.

A key idea in many reinforcement learning algorithms is that these values can be learned incrementally from experience using prediction errors. The earlier introduced Temporal-difference (TD) learning methods rely on comparing the predicted value of the current state to a bootstrap estimate based on the next state [Sutton and Barto, 2018]. For a state-value function, the TD error at time t is given by

$$\delta_t \;=\; r_t + \gamma V(s_{t+1}) - V(s_t),$$

where δ_t is the TD prediction error at time step t, r_t is the reward received after transitioning out of state s_t, $\gamma \in [0, 1]$ is a discount factor that determines how much future rewards are

downweighted relative to immediate ones, $V(s_{t+1})$ is the current estimate of the value of the next state s_{t+1}, and $V(s_t)$ is the current estimate of the value of state s_t. The value estimate is then updated according to

$$V(s_t) \leftarrow V(s_t) + \alpha\, \delta_t,$$

where $\alpha \in (0, 1]$ is a learning rate controlling the step size of each update. The prediction error δ_t quantifies how much better or worse the outcome was than expected, and the learning rule adjusts value estimates so that future predictions more closely match experienced rewards.

Within this framework, an important distinction is often drawn between *model-free* and *model-based* reinforcement learning [Dayan and Niv, 2008, Daw et al., 2011]. Model-free algorithms, such as standard TD learning, directly learn cached values from experience without representing the underlying transition structure $P(s' \mid s, a)$ of the world. They are computationally simple and can explain habitual, well-practised behaviours, but they tend to adapt slowly when contingencies change, because value estimates must be relearned incrementally [Dickinson and Balleine, 1994, Valentin et al., 2007]. Model-based algorithms, by contrast, learn or maintain an explicit internal model of the environment's dynamics and reward structure, and use this model to plan ahead by simulating possible future trajectories [Daw et al., 2011]. This allows flexible, goal-directed behaviour and rapid adaptation to changes in contingencies, but at the cost of greater computational complexity and reliance on an accurate internal model.

Reinforcement learning has been fruitfully applied to a wide range of animal and human learning tasks. In instrumental conditioning paradigms, model-free algorithms capture gradual acquisition of action–outcome contingencies and habitual responding, while model-based accounts explain phenomena such as sensitivity to outcome devaluation, revaluation, and contingency degradation [e.g., Dayan and Niv, 2008, Keramati et al., 2011, Lee et al., 2014]. In humans, model-derived value and prediction error signals have been shown to correlate with activity in striatal, prefrontal, and midbrain regions, supporting the idea that TD-like mechanisms are implemented in corticostriatal circuits and modulated by dopaminergic signals [e.g., O'Doherty et al., 2004, Daw et al., 2011, Niv et al., 2015].

Standard reinforcement learning models thus provide a well-grounded account of how values are updated from reward and punishment signals. However, their effectiveness depends entirely on the state space over which learning operates, a point developed in the next subsection.

1.3 The State Representation Problem

Standard reinforcement learning models typically begin from the assumption that the relevant *state space* has already been specified. That is, the agent is provided with a set of states that are presumed to capture all decision-relevant aspects of the environment, and learning proceeds by updating values [Niv, 2019]. In simple simulations the state space can be hand-engineered by

the modeller, and a well-chosen representation is often sufficient for RL algorithms to perform well. This convenience, however, sidesteps a fundamental question: where do the states come from? The approach breaks down in complex, real-world environments where the relevant state space is not obvious in advance, and choosing an appropriate representation becomes itself a central part of the learning problem [Radulescu et al., 2021].

A fixed, pre-specified representation has several important consequences. First, if states that are distinct in the representation nevertheless share a common underlying cause or reward source, standard value learning will treat them as independent and fail to exploit this shared structure. There is no mechanism for deciding that two nominally different states should share the same value, or for generalising from one to the other [e.g., Liu et al., 2021b, Wimmer et al., 2012]. Second, when the latent structure of the environment changes, for example, when previously correlated options become independent, or when new contingencies emerge, a rigid state space can lead to slow or inappropriate adaptation [e.g., Momennejad et al., 2017, Courville et al., 2006]. Values tied to an outdated partitioning of states must be relearned incrementally, even if a different grouping would permit much faster adjustment. Third, if the representation fails to distinguish states that differ in important ways (for instance, by aliasing together distinct contexts), learning can converge on compromised or unstable value estimates that do not support flexible behaviour [Gershman et al., 2015].

What, then, makes a good representation? Two principles stand out. First, effective representations should compress irrelevant variation: raw sensory inputs are high-dimensional and contain many features inconsequential for future outcomes, and only those aspects that matter for prediction and control need to be retained [e.g., Gershman and Niv, 2015, Abel et al., 2018, Botvinick et al., 2019]. Second, representations should group experiences by their consequences: situations that share the same reward contingencies and optimal actions can be treated as equivalent, enabling value updates in one to generalise immediately to the others, while situations requiring different actions or predicting different outcomes should be kept distinct [Li et al., 2006, Niv, 2019, Liu et al., 2021b, Wimmer et al., 2012]. By exploiting discovered structure, such representations enable flexible behaviours such as one-shot generalisation: when an agent realises that two options share a latent reward source, an observed change in the value of one should immediately imply a corresponding change in the value of the other, without requiring further direct experience [Liu et al., 2021b].

These considerations highlight that performance in structured environments depends not only on how values are updated, but also on how states are represented. Recent work in computational and cognitive neuroscience has therefore emphasised the importance of learned task-state representations and internal models that capture latent structure [e.g., Gershman et al., 2010, Niv, 2019, Behrens et al., 2018, Schuck et al., 2016]. The central question for the remainder of this chapter is how such representations are acquired and refined: how agents discover which situations belong together, how they reorganise these groupings when structure changes, and what computational and neural mechanisms support this process.

1.3.1 Learning Internal Task Representations

If an appropriate state space cannot be taken for granted, a key challenge is to understand how agents construct internal *task states* that summarise the information needed for prediction and control. Rather than treating each momentary sensory input as a distinct state, it is often useful to posit latent states that capture the underlying situation an agent believes itself to be in: which task is currently active, which rules apply, or which reward contingencies are in force. These latent task states can be thought of as internal variables that integrate over observations and history to provide a compact description of the decision-relevant aspects of the environment [e.g., Gershman et al., 2010, Collins and Frank, 2013, Niv, 2019, Schuck et al., 2016].

Constructing such task states is closely related to *state abstraction*: the process of grouping situations that are equivalent for prediction and action, while keeping distinct those that differ in their consequences [Sutton et al., 1999, Abel et al., 2018]. Computational models have proposed that agents infer latent contexts or causes to explain observed patterns and to decide when to reuse or create new task states [e.g., Gershman et al., 2010, Gershman and Niv, 2015, Niv, 2019]. Empirical studies suggest that prefrontal and medial temporal regions encode abstract task variables and context-specific state spaces that go beyond immediate sensory input [e.g., Schuck et al., 2016, Wilson et al., 2014, Vaidya and Badre, 2022]. These internal representations allow behaviour to generalise across conditions that share structure and to change rapidly when inferences about the current context shift.

The problem of discovering useful state representations has a long history beyond cognitive neuroscience. In classical RL, McCallum [1995] proposed the "utile distinctions" framework, in which an agent learns to split or merge state representations based on whether a distinction is useful for predicting future reward, offering an early computational account of how state spaces can be constructed from interaction rather than specified in advance. More recently, deep reinforcement learning has approached the problem from a different direction: by training large neural networks end-to-end on raw sensory input, useful representations emerge implicitly as a byproduct of optimising task performance [Botvinick et al., 2019]. The cognitive neuroscience perspective developed in this chapter complements these approaches by asking how biological agents solve the same problem, with particular attention to the representational formats and neural mechanisms involved.

From this perspective, a core part of the learning problem is to decide which experiences belong together and should be assigned to the same task state, and which should be kept separate. This assignment cannot be fixed a priori, but must itself be learned from experience, based on regularities in observations and outcomes [Gershman and Niv, 2010]. The following sections develop two complementary accounts of how such structured representations are constructed: cognitive maps, which describe the form such representations can take, and latent cause inference, which addresses how they are acquired.

1.4 Cognitive Maps as Structured Representations

1.4.1 From Scalar Values to Relational Maps

Standard value functions assign each state a summary of its expected reward consequences, but they do not represent how states relate to one another: which states are reachable from which, which share a common cause, or how the environment is organised more broadly. The notion of a *cognitive map*, introduced by Tolman [1948] in the context of spatial navigation, offers a richer representational format that captures precisely this relational structure. In a classic demonstration, Tolman [1948] showed that rats allowed to explore a complex maze without reward later took the shortest available path to a newly introduced goal, performing as well as animals that had been rewarded throughout training. This *latent learning* result implied that the animals had built an internal representation of the maze layout during unrewarded exploration, a representation that could be flexibly deployed once a goal was introduced. The finding was difficult to reconcile with stimulus–response accounts of learning, and motivated the idea that agents construct map-like internal models of their environment that exist independently of any particular reward contingency [O'Keefe and Nadel, 1978]. In modern computational terms, cognitive maps can be viewed as relational or graph-like internal models in which states are nodes and their relationships, such as transitions, distances, or shared latent causes, are encoded in the edges or geometry of the space [Behrens et al., 2018, Whittington et al., 2020, Stachenfeld et al., 2017, Schapiro et al., 2013]. In the framework of model-based RL, this structural knowledge is formalised as a transition matrix defining the probability of moving from one state to another given a specific action. The transition matrix can thus be seen as a computational instantiation of the cognitive map: it encodes the relational structure of the environment independently of reward, providing the backbone on which value computations operate.

The key distinction, then, is between *reward information*, what outcomes a state yields, and *structural information*, how states connect to and relate to one another. Separating these two kinds of knowledge has important functional consequences. Values can be learned and updated on top of a relatively stable structural backbone, enabling computations, such as planning, shortcut-finding, and flexible revaluation, that rely on understanding how different parts of the environment fit together [Stachenfeld et al., 2017, Behrens et al., 2018, Doll et al., 2012]. Structured, map-like representations support flexible planning by allowing agents to simulate multi-step trajectories and evaluate alternative routes [Pfeiffer and Foster, 2013, Doll et al., 2012]; facilitate generalisation across related states [Garvert et al., 2023]; and provide a natural substrate for integrating new information without relearning the entire value function. These properties have motivated proposals that cognitive maps form a core ingredient of flexible, model-based control in humans and other animals [Bellmund et al., 2018, Behrens et al., 2018, Bellmund et al., 2019b, Constantinescu et al., 2016].

1.4.2 Neural Substrates of Cognitive Maps

The idea that the brain constructs cognitive maps has its origins in spatial navigation. Place cells in the hippocampus fire selectively when an animal occupies a particular location [O'Keefe and Dostrovsky, 1971], while grid cells in medial entorhinal cortex exhibit periodic hexagonal firing patterns providing a metric representation of space [Hafting et al., 2005, Moser et al., 2008]. These are complemented by a wider family of spatially tuned cell types with presumed complementary functions, including head direction cells that encode the animal's facing direction [Taube et al., 1990], border cells that fire near environmental boundaries [Moser et al., 2008], and speed cells whose firing rate scales with running velocity [Kropff et al., 2015]. Together, these findings established the hippocampal–entorhinal system as a canonical substrate for spatial cognitive maps [Moser et al., 2017].

It is worth noting, however, that individual place cells represent locations *within* a map rather than the map's structure itself. A single place cell signals where the animal is, but does not directly encode how that location connects to others. The relational structure of the map, the topology and metric distances between locations, is better understood as an emergent property of the population-level activity patterns across these diverse cell types, with grid cells in particular proposed to provide a coordinate framework that captures distances and directions between locations [Hafting et al., 2005, Moser et al., 2008]. More recent work has extended this mapping role to abstract and task-related domains. Human fMRI studies show that hippocampal and entorhinal activity patterns encode positions in spaces defined by non-spatial features [Constantinescu et al., 2016, Garvert et al., 2023, 2017], and that OFC and medial prefrontal regions represent task-state structure and relationships [Wilson et al., 2014, Schuck et al., 2016, Wikenheiser and Schoenbaum, 2016]. The role of OFC is further supported by animal work: recordings in rodent OFC reveal neurons that encode latent task states, including hidden variables not directly observable in the current stimulus, and these representations evolve as animals learn the structure of a task [Zhou et al., 2021, Zong et al., 2025]. Recent work has argued that OFC and ventromedial prefrontal cortex maintain representational spaces that encode both task-state structure and value, with the two kinds of information represented in partially overlapping but dissociable formats [Moneta et al., 2024]. Representational similarity analyses confirm that activity in hippocampal and prefrontal regions is organised according to graph structure, transition probabilities, or learned relational structure, even when reward is held constant [Stachenfeld et al., 2017, Behrens et al., 2018, Deuker et al., 2016, Bellmund et al., 2019b], suggesting that these circuits encode structural knowledge about the environment rather than reward value per se.

Taken together, these findings suggest that the brain may not maintain a single, unified cognitive map but rather multiple, partially overlapping maps distributed across hippocampal, entorhinal, and prefrontal regions, each capturing different aspects of environmental structure [Behrens et al., 2018, Moneta et al., 2024]. By separating relational structure from value as-

signment, these neural maps are hypothesised to support model-based computations in which values are derived from an underlying structural model and can be updated efficiently when contingencies change [Stachenfeld et al., 2017, Momennejad et al., 2017, Russek et al., 2017].

1.4.3 Cognitive Maps Beyond Physical Space

Hippocampal-prefrontal circuits appear capable of representing positions within spaces defined by non-spatial features, stimulus dimensions, or latent task variables [Behrens et al., 2018]. Key evidence for this view comes from animal as well as human work. In rodents, hippocampal neurons have been found to encode elapsed time during a delay period in which the animal runs in place, forming sequential "time cell" representations that tile the temporal interval much as place cells tile spatial locations [MacDonald et al., 2011, Kraus et al., 2013]. More strikingly, Aronov et al. [2017] showed that when rats performed a non-navigational task in which they manipulated sound frequency to obtain reward, hippocampal and entorhinal neurons exhibited tuning curves along the frequency dimension that closely resembled the spatial tuning of place and grid cells, demonstrating that map-like coding generalises to entirely non-spatial dimensions even in rodents.

In humans, fMRI studies have shown grid-like coding in entorhinal cortex for abstract stimulus dimensions [Constantinescu et al., 2016, Doeller et al., 2010], and hippocampal activity patterns that reflect positions along learned graphs or manifolds defined by stimulus relationships [Garvert et al., 2017, 2023, Deuker et al., 2016, Bellmund et al., 2019b]. The same representational principles extend to social cognition: Tavares et al. [2015] found that hippocampal activity tracked participants' positions within a two-dimensional social space defined by power and affiliation, suggesting that the hippocampus maps social relationships using the same geometric framework it applies to physical space.

Together, these converging lines of evidence from animal electrophysiology, human neuroimaging, and diverse task domains motivate a broader view of cognitive maps as internal models defined over task states or latent dimensions rather than physical locations. States that share similar consequences, transition structure, or latent causes are represented as nearby; states with different roles are more distant [Stachenfeld et al., 2017, Behrens et al., 2018]. In the task studied in this thesis, for example, sequences sharing a reward source occupy neighbouring positions in a latent reward space, such that a value change for one should propagate to structurally nearby sequences but not to independent ones. A prediction that depends entirely on the agent having acquired an appropriate internal representation.

Crucially, such abstract maps must be learned from experience rather than assumed. The next section addresses how this acquisition can occur through latent cause inference.

1.5 Structure Learning as Latent Cause Inference

1.5.1 From Given Maps to Learned representations

The cognitive map framework offers a compelling account of how relational structure can be represented in the brain, but leaves open the question of how such structure is acquired in the first place. Many theoretical treatments sidestep this by assuming the map is given, specifying topology or graph structure by design such that learning reduces to estimating values on a fixed pre-learned or given structure [e.g., Liu et al., 2021b].

In natural settings, by contrast, the map itself is rarely given. Agents must infer which states are connected, which options share outcomes or latent causes, and how different parts of the environment are organised. From this perspective, *structure learning* refers to the process of discovering the latent organisation of states and contingencies from experience [Gershman and Niv, 2010]. This includes inferring which stimuli belong to the same context, which actions lead to similar consequences, and which states can be grouped together because they share a hidden reward source or rule [e.g., Gershman et al., 2010, 2015, Collins and Frank, 2013].

A simple example is context-dependent learning, in which the same stimulus–response pair can have different outcomes in different contexts (such as environments, task phases, or latent modes of a task). To perform well, an agent must infer when a change in context has occurred and adjust its internal state accordingly [Gershman et al., 2010, Collins and Frank, 2013]. Behavioural phenomena such as rapid switching between different policies, renewal effects after context changes, or sudden insight when a hidden structure is discovered suggest that humans and animals infer latent task states that organise experience into distinct contexts [e.g., Wilson and Niv, 2012, Gallistel and Gibbon, 2000]. Similarly, in tasks with correlated options, performance often shows abrupt improvements once participants recognise that certain options share a common cause, indicating a restructuring of the internal representation rather than a gradual, option-by-option relearning. The remainder of this section develops structure learning more formally in terms of latent cause inference, in which agents are assumed to infer hidden causes that generate their observations and to use these inferred causes as the building blocks of internal cognitive maps.

1.5.2 Latent Cause Inference as a Framework

One influential way to formalise structure learning is in terms of *latent cause inference* (LCI). LCI draws on a long tradition of Bayesian latent variable modelling in statistics and machine learning. The core premise of latent variable models is that observed data are generated by hidden (unobserved) variables, and that learning consists of recovering these hidden variables from the pattern of observations [Bishop and Nasrabadi, 2006]. This is a general and powerful framework: mixture models explain clusters in data by positing that each observation was

generated by one of several hidden sources [Dempster et al., 1977]; hidden Markov models posit that a sequence of observations is driven by transitions among unobserved states; and topic models decompose documents into mixtures of latent themes [Blei et al., 2003]. In each case, the structure of the data is explained by inferring which latent variable was responsible for which observation.

LCI applies this logic to the problem of learning from experience. Here, the latent variables are interpreted as hidden "causes" or "contexts" that generate the patterns of observations and outcomes an agent encounters [e.g., Gershman et al., 2010, 2015]. Each latent cause corresponds to a distinct configuration of contingencies or rules, for example, a particular task context, reward structure, or environmental mode, and sequences of observations are explained by mapping them to their underlying causes.

The core computational idea is to cluster observations into latent causes based on their similarity and predictive success. Observations that tend to co-occur, share similar outcome statistics, or are better predicted when grouped together are inferred to arise from the same hidden cause [Gershman and Niv, 2010]. Conversely, when prediction errors become systematically large, this signals that the current set of causes may be insufficient, prompting the creation of a new cause to better account for the data [Gershman and Niv, 2015]. In this way, the agent maintains a set of inferred causes and an assignment of past observations to these causes, which jointly define an internal, structured representation of the task.

Bayesian formulations of LCI make these ideas precise by specifying a prior over possible cause structures and updating this prior in light of new evidence. Nonparametric approaches, such as those based on the Chinese Restaurant Process (CRP) or related Dirichlet process priors, offer a particularly flexible class of models [Rasmussen, 1999, Sanborn et al., 2010, Griffiths and Ghahramani, 2011]. In such models, the number of latent causes is not fixed in advance. Instead, the prior favours reuse of existing causes that already explain many observations, while still allowing new causes to be created when warranted by the data. The probability that a new observation is assigned to an existing cause increases with how frequently that cause has been used and how well it predicts the observation; the probability of assigning it to a new cause depends on a concentration parameter that controls the model's tendency to create new clusters [Gershman et al., 2010].

Together, these properties allow LCI models to balance structural reuse against flexible expansion of the internal model. In the context of value-based decision-making, the inferred latent causes can be viewed as the building blocks of internal cognitive maps, determining which states share value information and how changes in one part of the environment propagate to others.

1.6 Non-Local Value Updating Through Latent Cause Inference

1.6.1 Value Propagation via Shared Latent Causes

When observations corresponding to different nominal states are assigned to the same latent cause, they come to share a common set of parameters governing their expected outcomes. Rather than learning separate values for each state in isolation, the agent effectively learns about the value of latent causes, and individual states derive their value from their cause membership [e.g., Gershman et al., 2010, Gershman and Niv, 2015, Pettine et al., 2021, Mirea et al., 2024]. This organisation supports *non-local* value updating: when a new outcome is observed for one state, the resulting prediction error updates beliefs about the underlying latent cause, and because other states associated with that cause depend on the same latent parameters, their inferred values change as well, even if they have not been directly experienced in the current context [Liu et al., 2021b, Lehnert et al., 2020].

The ability to create, merge, or split latent causes also provides a route to flexible adaptation when the underlying structure changes. If outcomes that were previously well explained by a shared cause begin to diverge systematically, prediction errors will increase, raising the posterior probability that a new cause should be introduced or that observations should be reassigned [Courville et al., 2006, Pisupati et al., 2023]. This leads to a restructuring of the internal representation: states formerly grouped together may be separated into distinct causes, allowing value estimates to diverge, while states that begin to share similar outcome statistics may be merged [Collins and Frank, 2013]. Because the number and granularity of inferred causes depend on factors such as prior expectations and sensitivity to prediction errors, latent cause models naturally predict individual differences in generalisation. Learners who infer fewer, broader causes will generalise more strongly, facilitating transfer but risking overgeneralisation; learners who infer narrower causes will show more cautious generalisation and may adapt more readily when contingencies change, but may fail to exploit shared structure efficiently [e.g., Gershman et al., 2010, Collins and Frank, 2013].

1.6.2 Predictive and Successor-Like Representations

Latent cause inference supports generalisation by clustering states that share hidden causes, so that value updates to one state propagate to others via their common cause membership. A distinct route to generalisation comes from *predictive representations*: internal codes that embed each state according to the future states or features it predicts, rather than according to inferred latent variables [e.g., Sutton et al., 1999, Gershman, 2018]. States that lead to similar successors are thereby represented similarly, and value information can propagate across states linked by shared predictive structure, even without any explicit inference over hidden causes.

A prominent example is the *successor representation* (SR) [Dayan, 1993, Momennejad et al., 2017, Russek et al., 2017, 2021], which encodes for each state the expected discounted occupancy of all future states under a given policy. Importantly, the SR is typically learned incrementally through temporal-difference updates, much like scalar value functions: at each transition, the agent updates its estimate of which future states will be visited from the current state, using a prediction error over state occupancies rather than over rewards [Gershman, 2018]. This makes the SR computationally analogous to model-free value learning, but operating over state predictions rather than reward predictions. The resulting factorisation separates knowledge about the predictive structure of the environment from knowledge about current reward assignments. When rewards change but transition structure remains stable, values can be rapidly recomputed by combining the existing SR with the new reward function, without relearning the map from scratch.

However, the SR's utility for flexible replanning is constrained by its dependence on the policy under which it was learned. Because it encodes expected future trajectories under a specific policy, it supports accurate value computation only when the agent continues to act in a similar way. When goals shift and the optimal policy changes accordingly, the stored SR no longer correctly predicts which states will be visited, rendering it unreliable precisely when flexible replanning is most needed [Russek et al., 2017, Piray and Daw, 2021]. The *default representation* (DR) proposed by Piray and Daw [2021] addresses this limitation by anchoring the predictive map to a fixed *default policy*: a stable set of baseline behavioural assumptions, rather than to the currently optimal one. Because the DR does not track the agent's evolving choice policy, it remains a stable substrate for planning even as goals and reward contingencies change. Deviations from the default are permitted and optimised over, but at a cost explicitly traded off against expected reward.

These predictive approaches and latent cause inference thus offer complementary accounts of how value information can propagate beyond directly experienced states. Both achieve generalisation by establishing learned associations across states that allow value updates to spread beyond the directly observed option. They differ, however, in the nature of these associations. In LCI, states become linked through their inferred membership in a shared latent cause, such that a value change for one state updates the cause and thereby affects all other states assigned to it. In the SR and DR, states become linked through incrementally learned predictions about future state occupancies, such that value propagates along the predictive map connecting a state to its expected successors. Both require that the agent maintain internal structure linking states together, but they differ in how that structure is acquired and in the kinds of generalisation they most naturally support.

1.7 Empirical Project on Learning Latent Structure and Non-Local Value Propagation

The preceding sections have developed a computational account of how agents discover the latent structure of their environment through latent cause inference and use that structure to propagate value information across states linked by shared causes or predictive associations. A natural question is how such non-local value updates are implemented in concrete learning systems. In reinforcement learning, Sutton [1991] proposed DYNA, an architecture in which an agent not only learns from real interactions with the environment but also maintains an internal model of it and uses that model to generate simulated experiences. Between real steps, the agent "replays" transitions sampled from its model and updates value estimates from these imagined experiences, effectively learning offline without further environmental interaction [Sutton and Barto, 2018]. Prioritised variants of this scheme sample transitions in proportion to their expected utility, for example based on prediction error magnitude, directing offline computation toward the parts of the state space where updating would be most beneficial [Mattar and Daw, 2018]. In biological systems, hippocampal replay has been proposed as a candidate neural mechanism for such model-based offline updating, a possibility developed in detail in Chapter 2.

Prior work has proposed that structured internal representations supporting flexible generalisation can be understood as cognitive maps [Tolman, 1948, Behrens et al., 2018], and that their acquisition can be described in terms of latent cause inference [Collins and Frank, 2013, Lloyd and Leslie, 2013, Gershman and Niv, 2015]. In this context, replay could serve as the mechanism by which value information is propagated across states that share an inferred latent cause, implementing non-local updating through offline reactivation of cause-relevant experiences. However, existing studies have not directly linked latent cause inference to replay-based mechanisms, leaving open whether replay shapes the acquisition of latent structure, exploits it for non-local value updating, or both.

The study presented in Chapter 5 was designed to address this gap. In this study, 52 participants performed a sequential bandit task with hidden reward correlations over two days. On Day 1, participants encountered a novel environment in which three of four bandits shared a latent reward source, while the fourth had an independent reward. Crucially, participants were never instructed about this correlation structure and therefore had to infer it from experience. Reward means drifted and reversed over time, requiring continual adaptation and providing a test of whether participants could generalise value changes to unobserved but structurally linked options. On Day 2, the latent reward structure was covertly altered mid-session, probing how flexibly participants could reorganise their internal representations when the underlying contingencies changed.

A central aim of the study was to test whether a latent cause inference model, grounded

in the principles of structure learning outlined in the preceding sections, provides a better account of participants' behaviour than classical temporal-difference models with and without hard-coded generalisation. The task was designed such that successful performance required discovering which bandits share a common reward source and using this inferred structure for non-local credit assignment, precisely the computational problem that LCI models are built to solve. In addition to these behavioural and computational questions, the study combined fMRI-based measures of neural replay and representational change with the computational model, providing a neural instantiation of the non-local updating mechanisms discussed above. These replay-related aspects are introduced and discussed in detail in Chapter 2.

Chapter 2

Neural Replay and Offline Updating

The previous chapter argued that flexible generalization in value-based decision-making depends on internal representations that capture latent task structure and support non-local value updating. The central question of the present chapter is how such offline, structure-based updates might be implemented in the brain. A prominent candidate mechanism is *neural replay*: the rapid, sequential reactivation of activity patterns corresponding to past or hypothetical states during rest, quiet wakefulness, or sleep [Wittkuhn et al., 2021]. Replay has been proposed to provide a neural substrate for reprocessing experience offline, enabling prediction errors to update value information without direct environmental input [e.g., Foster, 2017, Liu et al., 2021b].

Replay was first observed in the rodent hippocampus. Seminal recordings showed that place-cell ensembles active during spatial navigation were reactivated during subsequent sleep in the same sequential order but at a highly compressed timescale [Wilson and McNaughton, 1994, Skaggs and McNaughton, 1996]. Subsequent work revealed that similar sequence reactivation occurs during quiet wakefulness, frequently coupled to hippocampal sharp-wave ripple (SWR) events, and can represent both recently experienced and more remote trajectories [e.g., Davidson et al., 2009, Foster, 2017, Ólafsdóttir et al., 2018, Gupta et al., 2010]. Crucially, disrupting SWRs during post-learning sleep impairs subsequent memory performance, providing causal evidence that ripple-associated replay contributes to memory consolidation [Girardeau et al., 2009]. Over time, the concept of replay has broadened well beyond these original spatial sequences to encompass non-spatial, task-related, and abstract state trajectories in both animals and humans.

Replay-like phenomena have now been documented across species, brain regions, and representational domains. In rodents, hippocampal replay is closely coordinated with activity in medial prefrontal and sensory cortices, pointing to a systems-level mechanism for memory redistribution and decision-making [Jadhav et al., 2016, Ólafsdóttir et al., 2016, 2018]. Ji and Wilson [2007] demonstrated simultaneous replay in the hippocampus and visual cortex during slow-wave sleep, with cortical sequences lagging hippocampal sequences by a brief interval consistent with directed hippocampal-to-cortical information transfer. Disrupting hippocampal–prefrontal coordination during SWR events impairs spatial memory, further

underscoring the functional importance of distributed replay [Jadhav et al., 2016]. In humans, multivariate decoding of MEG and fMRI signals has revealed temporally structured reactivation of stimulus and task representations during rest, brief pauses, and reward processing, often reflecting internal models or task rules rather than mere sensory order [e.g., Liu et al., 2019, 2021b, Schuck and Niv, 2019]. Together, these findings establish replay as a widespread phenomenon operating across hippocampal–cortical networks in the service of learning, planning, and memory.

This chapter pursues four aims. First, it clarifies what is meant by "replay" in the context of this thesis, distinguishing it from the broader notion of reactivation, and outlines the methods available for detecting replay in humans. Second, it reviews the major empirical findings on replay in animals and humans, focusing on those most relevant to value-based learning and the empirical work of this thesis. Third, it connects these findings to computational theories of learning and consolidation, including offline reinforcement learning, and memory consolidation that provide a normative framework for understanding when and what to replay [e.g., Sutton and Barto, 2018, Mattar and Daw, 2018, Wittkuhn et al., 2021, Kirkpatrick et al., 2017]. Finally, it uses this framework to motivate the specific hypotheses tested in the present experiments: that replay operates over inferred latent reward structure, implements prioritised nonlocal value updates, and tracks changes in internal representations as the underlying structure of the environment changes.

2.1 What Is Replay? From Reactivation to Structured Sequences

2.1.1 Reactivation vs. Replay

Discussions of replay often begin from a broader notion of *reactivation*: the reinstatement of neural activity patterns associated with previous experiences, regardless of whether this reinstatement is temporally structured [Tingley and Peyrache, 2020, van Der Meer and Bendor, 2025]. Reactivation is typically quantified via pattern similarity or decoding approaches, in which a neural pattern observed during rest, sleep, or a later task phase is compared to patterns elicited during initial encoding; increased similarity is taken as evidence that the earlier representation has been re-engaged [e.g., Skaggs and McNaughton, 1996, Chen and Wilson, 2023]. This definition is agnostic about temporal order: it focuses on *what* is being reinstated and *where* in the brain this occurs.

The term *replay*, by contrast, is reserved for temporally *structured* reactivation, i.e., sequences of neural states whose ordering reflects or systematically transforms prior experience. In classical rodent work, replay denotes coordinated place-cell firing sequences that recapitulate spatial trajectories at a compressed timescale during sleep or quiet wakefulness [Wilson

and McNaughton, 1994, Skaggs and McNaughton, 1996, Foster, 2017]. More broadly, replay can be defined as the sequential reactivation of a series of states with characteristic properties such as temporal compression, directionality (forward or reverse), and intermittency [Ólafsdóttir et al., 2018, Chen and Wilson, 2023]. Replay, in this sense, is about *how* past states unfold over time, not just which states are active.

Recent conceptual work has emphasised the importance of maintaining this distinction. Chen and Wilson [2023] argue that phenomena labelled "replay" differ along several key dimensions, including the nature of the represented content (spatial vs. non-spatial), temporal structure (continuous vs. fragmented, compressed vs. veridical time), and brain state (sleep vs. awake), and that conflating them obscures both mechanistic understanding and cross-study comparison. Similarly, Wittkuhn et al. [2021] highlight that reactivation without clear sequential organisation should not automatically be interpreted as replay, and that a careful distinction between content ("what/where") and dynamics ("how": speed, direction, temporal warping) is essential when comparing results across species and methods.

In this thesis, *replay* refers to phenomena showing statistically reliable temporally ordered reactivation of task states (for instance, backward sequences following reward receipt).

2.1.2 Key Dimensions

Replay events vary along several dimensions that organise empirical findings and link them to candidate computational roles [Foster, 2017, Ólafsdóttir et al., 2018, Chen and Wilson, 2023, van Der Meer and Bendor, 2025]. Four dimensions are particularly relevant here.

Direction and locality

A first dimension concerns replay *direction*. *Forward* replay, in which sequences unfold in the same order as during behaviour, has been interpreted as simulating prospective trajectories or re-experiencing recent paths. *Reverse* replay, by contrast, runs backward from a goal location toward preceding states and is especially prominent immediately after reward delivery, consistent with backward value propagation for credit assignment [Foster, 2017]. The strength of reverse replay is modulated by reward magnitude, pointing to a direct functional link with value learning rather than a passive echo of recent experience [Foster, 2017]. Forward replay at choice points and prior to trajectory initiation, in turn, is consistent with a prospective planning function [Pfeiffer and Foster, 2013].

A related distinction concerns *locality*. Local replay begins near the animal's current position, whereas remote replay represents locations far from the current state, including unvisited paths and putative shortcuts [e.g., Gupta et al., 2010, Karlsson and Frank, 2009]. Remote replay during quiet wakefulness indicates the accesses to a broad internal model of the environment rather than being spatially anchored to recent experience [Karlsson and Frank, 2009]. Together,

these variations demonstrate that replay flexibly samples trajectories relevant for different computational purposes, rather than merely recapitulating the most recent experience.

Temporal properties

A second dimension concerns the temporal relationship between replay and the original experience. Classical hippocampal replay unfolds at a much faster timescale than behaviour, with compression factors of approximately 5–20× relative to the original traversal [Skaggs and McNaughton, 1996, Nádasdy et al., 1999, Lee and Wilson, 2002]. However, replay trajectories can also exhibit variable speed, pauses, and discontinuous jumps rather than smooth linear compressions [Gupta et al., 2010, Chen and Wilson, 2023]. Sequences sometimes stitch together segments from different episodes or explore routes not previously experienced as a single continuous path. These properties indicate that replay can be a generative process, recombining elements of experience to construct novel trajectories, rather than merely accelerating veridical memories [Kurth-Nelson et al., 2023].

Content: what is replayed

A third dimension concerns the state space over which replay sequences are defined. Classical work focused on spatial trajectories encoded by hippocampal place cells. More recent studies have demonstrated replay of non-spatial and task-related states, including abstract positions in graph-like environments, latent task states, stimulus categories, and task rules [e.g., Schuck et al., 2016, Liu et al., 2019, Kurth-Nelson et al., 2023]. In humans, multivariate decoding has revealed sequential reactivation of representations corresponding to task states, decision options, or conceptual features, sometimes in orders that reflect internal cognitive structure rather than overt sensory sequences [Liu et al., 2019, Schwartenbeck et al., 2023, Wittkuhn et al., 2025]. This broader view aligns replay with the cognitive map framework [Behrens et al., 2018], in which the relevant "space" can be spatial, conceptual, or defined by latent task structure.

Brain state: when replay occurs

A fourth dimension concerns the global brain state during replay. Replay was initially characterised during slow-wave sleep, where SWR-associated sequences have been linked to systems consolidation and long-term memory [Foster, 2017, Ólafsdóttir et al., 2018]. Subsequent work identified robust replay during quiet wakefulness and brief pauses in behaviour, often following reward delivery or at decision points [Foster, 2017, van Der Meer and Bendor, 2025, Liu et al., 2021b]. van Der Meer and Bendor [2025] characterise awake replay as the brain being "off the clock but on the job": although the animal is not actively navigating, replay performs targeted value backups and tags experiences for preferential consolidation during

subsequent sleep, positioning awake replay as computationally distinct from sleep replay in its emphasis on rapid, experience-proximal updating.

Growing evidence extends replay-like phenomena to internally oriented cognition in humans, including mind-wandering, rest, and brief task pauses, when external demands are minimal [Wittkuhn et al., 2021, Chen and Wilson, 2023, Wittkuhn et al., 2025]. Most recently, replay has been observed during active task performance, where it may support ongoing deliberation by assembling candidate solutions during compositional inference [Schwartenbeck et al., 2023], though this "on-task" replay remains less well established than its offline counterpart. Different brain states may thus emphasise different functions: long-term consolidation during sleep, shorter-term credit assignment or planning during awake rest, and real-time hypothesis testing during ongoing cognition.

These dimensions (direction and locality, temporal properties, state space, and brain state) provide a scaffold for thinking about replay as a heterogeneous family of processes rather than a single mechanism. The following sections use them to organise empirical findings and to map different aspects of replay onto candidate computational roles: systems consolidation, offline credit assignment, planning, and structure learning.

2.1.3 Computational Functions of Replay

The phenomenological diversity of replay, spanning variation in direction, timing, content, and brain state, suggests that it serves multiple computational functions rather than a single role. Multiple functional roles have been proposed, each grounded in a different computational framework and supported by partially overlapping empirical evidence.

Memory consolidation

On this view, replay during sleep gradually transfers information from the hippocampus to the neocortex, stabilising memories and integrating them with pre-existing knowledge structures. This account aligns with the complementary learning systems (CLS) framework [McClelland et al., 1995] and the active systems consolidation hypothesis [Born et al., 2006, Diekelmann and Born, 2010, Klinzing et al., 2019], in which SWR-associated replay provides the vehicle for hippocampal-to-cortical dialogue during slow-wave sleep. Supporting evidence includes the temporal coupling of replay with sleep oscillations, memory impairment following SWR disruption [Girardeau et al., 2009], and coordinated hippocampal–cortical replay during sleep in which cortical sequences lag hippocampal sequences, consistent with directed information transfer [Ji and Wilson, 2007]. Furthermore, the strength of post-learning replay predicts subsequent memory performance: Dupret et al. [2010] showed that reactivation of place-cell ensembles during post-learning rest predicted the stabilisation of newly formed spatial representations on the following day, directly linking replay to behavioural outcomes. This consolidation function is examined in detail in Chapter 3.

Offline reinforcement learning and credit assignment

From this perspective, replay implements experience replay analogous to DYNA-style algorithms [Sutton, 1991, Sutton and Barto, 2018], in which internally generated state transitions update value estimates without further environmental interaction. Reverse replay from reward locations is particularly consistent with this account: backward propagation of reward information along experienced trajectories provides an efficient credit assignment mechanism, and the strength of reverse sequences scales with reward magnitude [Foster, 2017]. Prioritised replay schemes further predict that replay content should be biased toward transitions with large prediction errors or high expected learning utility [Mattar and Daw, 2018]. Human MEG data provide direct support: Liu et al. [2021b] showed that backward replay following reward receipt preferentially represented non-local paths, that its strength scaled with a normative utility measure combining need and gain, and that stronger replay predicted more efficient non-local value learning at the trial-by-trial level.

Planning and prospective evaluation

Forward replay at decision points has been interpreted as mental simulation in which the agent evaluates candidate trajectories before committing to an action [Pfeiffer and Foster, 2013]. Supporting this account, forward replay at choice points predicts the animal's subsequent navigational decision. More broadly, model-based planning requires the ability to simulate multi-step trajectories through an internal model, and replay provides a candidate neural mechanism for such simulation [Mattar and Daw, 2018, Kurth-Nelson et al., 2023]. Recent work has extended this idea to compositional planning, in which replay assembles novel sequences from familiar elements to solve problems not directly experienced. Schwartenbeck et al. [2023] demonstrated that during a compositional construction task, human MEG replay sequences progressively converged on the correct configuration of building blocks, consistent with a generative hypothesis-testing process that evaluates candidate solutions through sequential sampling.

Structure learning and map construction

On this view, replay does not merely exploit an existing cognitive map but actively contributes to building and refining it. Evidence comes from observations that replay often represents trajectories the animal has not recently traversed, including remote locations and never-experienced shortcuts [Gupta et al., 2010], and that replay can lag behind behavioural learning rather than driving it [Gillespie et al., 2021]. Replay also shows a bias toward behaviourally relevant locations such as reward sites, suggesting that it samples from a structured internal model rather than passively echoing recent experience [Gupta et al., 2010]. These findings are difficult to reconcile with a purely evaluative or credit-assignment account and instead suggest

that replay stitches together separate experiences to form a coherent relational model [Ólafsdóttir et al., 2018, Kurth-Nelson et al., 2023]. Consistent with this idea, Wittkuhn et al. [2025] recently demonstrated backward sequential replay in human visual cortex during brief task pauses in a statistical learning paradigm without reward, with replay strength correlating with behavioural indices of successor representation learning, directly linking on-task replay to the acquisition of predictive cognitive maps. Recent theoretical work proposes that map-building and planning functions may operate as complementary processes, with replay prioritising one or the other depending on the agent's current uncertainty about environmental structure versus goals [Sagiv et al., 2025].

These four functions are not mutually exclusive. A single replay event may simultaneously contribute to consolidation, credit assignment, and structural refinement, and different brain states or task demands may shift the balance among them. The challenge for empirical work, taken up in the subsequent sections and in the empirical chapters of this thesis, is to identify experimental signatures that distinguish among these accounts.

2.2 Methods for Detecting Replay (in Humans)

2.2.1 Single-Unit and Population Decoding in Animals

The study of hippocampal replay originated in rodent electrophysiology, where simultaneous recordings from large populations of place cells enabled direct observation of compressed sequential reactivation during sleep and quiet wakefulness [Wilson and McNaughton, 1994, Skaggs and McNaughton, 1996, Lee and Wilson, 2002]. Following spatial navigation, the same spatially tuned neurons reactivate in temporally compressed bursts ($\sim$100 ms), recapitulating the sequential firing patterns observed during behaviour [Nádasdy et al., 1999, Davidson et al., 2009]. Candidate replay events are typically identified within sharp-wave ripples (SWRs), brief, high-frequency (150–250 Hz) oscillatory events in the hippocampal local field potential associated with synchronous population bursts [Buzsáki, 2015]. Replay is then quantified using decoding methods that estimate the posterior probability of position given ensemble spiking activity [Zhang et al., 1998] or weighted correlation between decoded position and time [Grosmark and Buzsáki, 2016]. Statistical significance is assessed against surrogate distributions generated by shuffling cell identities or spike timing within events [Ólafsdóttir et al., 2016, Tingley and Peyrache, 2020].

2.2.2 Multivariate Decoding and Sequence Detection in Humans

Investigating replay in the human brain poses fundamentally different methodological challenges. Non-invasive neuroimaging methods offer either high temporal resolution (MEG/EEG, $\sim$1 ms) or high spatial resolution (fMRI, $\sim$2 mm), but not both simultaneously. Moreover,

these methods record aggregate signals from large neuronal populations rather than individual spikes, precluding direct application of rodent replay analysis methods. In the case of fMRI, the measured signal reflects changes in blood oxygenation, capturing neural activity only indirectly through its metabolic consequences, with a temporal resolution on the order of seconds. Over the past decade, however, innovative analytical frameworks combining multivariate pattern analysis (MVPA) with sequence-sensitive statistical models have enabled the detection of replay-like phenomena in human neuroimaging data [Liu et al., 2022, Schuck and Niv, 2019, Wittkuhn and Schuck, 2021].

2.2.2.1 Classifier-based approaches

Current approaches to human replay detection share a common two-stage logic. First, multivariate pattern classifiers are trained on neuroimaging data acquired during task performance, when the identity of the neural representation is established by the experimental context (e.g., which stimulus is being viewed or which task state is active). Second, the trained classifiers are applied to data from periods of interest (rest, reward processing, or decision moments) to decode the time course of task-related neural representations [Norman et al., 2006, Tambini and Davachi, 2019]. This yields a time-by-state matrix of decoded reactivation probabilities, which forms the basis for subsequent sequence analyses. The critical methodological question is how to detect temporal regularities in these decoded time courses that reflect genuine sequential replay rather than noise or confounds.

2.2.2.2 Temporally delayed linear modelling (TDLM)

Liu et al. [2021a] developed temporally delayed linear modelling (TDLM), a comprehensive framework for detecting sequential structure in decoded neural representations. TDLM is particularly well suited for MEG and EEG data, where millisecond temporal resolution enables the detection of fast replay events, and has become one of the primary methods for human replay analysis. TDLM operates through a two-level general linear modelling (GLM) procedure. At the first level, for each time lag, a multiple linear regression estimates an empirical transition matrix, where each entry quantifies the degree to which the decoded representation of one state at a given time point predicts the representation of another state shortly afterwards, while controlling for all other states at that same time point. At the second level, this empirical transition matrix is projected onto a hypothesised task-related transition matrix to yield a single scalar measure of *sequenceness*. Repeating this procedure across a range of lags produces a sequenceness-by-lag function whose peak reveals the characteristic speed of replay events [Liu et al., 2019]. TDLM has been applied successfully to MEG data to reveal replay of both sensory and abstract task representations [Liu et al., 2019, Wimmer et al., 2020], to link backward replay to non-local reinforcement learning [Liu et al., 2021b].

2.2.2.3 sequential slope ordering analysis for fMRI

A complementary approach has been developed for fMRI, which offers superior spatial resolution but acquires data at a much slower rate (typically one volume every 1–3 s). Schuck and Niv [2019] demonstrated that despite these temporal constraints, sequential replay can be detected in human hippocampal fMRI signals by exploiting the prolonged hemodynamic response, which translates brief neural events into sustained BOLD signal changes.

In this approach, a multivariate classifier is trained to discriminate between distinct task states using fMRI patterns during task performance, then applied to resting-state data to produce a sequence of decoded state labels at each TR. The key insight is that if compressed replay events occur during rest, the sluggish BOLD response will "smear" the fast neural sequence across successive fMRI volumes, producing a characteristic ordering in decoded states [Schuck and Niv, 2019, Wittkuhn and Schuck, 2021].

The decoded sequences are summarised as an empirical transition matrix tabulating the frequency with which each pair of task states is decoded in consecutive volumes. This empirical structure is then compared to expected sequential distances in the task's state space. Schuck and Niv [2019] found that hippocampal resting-state patterns decoded after task performance exhibited significantly greater sequentiality than those decoded before the task, and that this sequentiality specifically reflected the abstract task-state structure rather than simpler sensory or attentional sequences.

This fMRI-based approach was further developed by Wittkuhn and Schuck [2021], who demonstrated that multivariate fMRI pattern analysis can detect sub-second neural activation sequences with high spatial precision. Using probabilistic classifiers trained on individual stimulus presentations, they showed that classifier-evidence dynamics within and across successive TRs systematically reflected the order and timing of fast image sequences presented at 32–128 ms between items. Critically, they also observed fast sequential replay in visual cortex during post-task rest, even when the task did not require memorisation, extending fMRI-based replay detection beyond the hippocampus and demonstrating sensitivity to behaviourally relevant timescales. This work further showed that probabilistic classifier evidence provides substantial statistical improvements over analyses based on the single most-likely decoded category [Schuck and Niv, 2019], and introduced a spectral modelling approach that enables inference about the speed of underlying neural sequences from the frequency content of fMRI sequentiality metrics.

Compared to TDLM, the fMRI-based approach trades temporal precision for spatial specificity. While fMRI cannot resolve the exact speed or directionality of replay with the same precision as MEG, it enables localisation of replay signals to specific brain regions, a capability particularly valuable given emerging evidence that replay occurs in a distributed, coordinated fashion across hippocampal and cortical circuits.

2.3 Replay, Latent Structure, and the Present Thesis

A recurring theme across the literature reviewed above is a tension between two broad accounts of replay's primary role. On one hand, replay may mainly propagate value information across an existing internal model: replay of unobserved paths scales with the degree to which value needs to propagate across a learned structure [Liu et al., 2021b, Pfeiffer and Foster, 2013, Mattar and Daw, 2018]. On the other hand, replay may actively contribute to building and refining that model: it often represents remote locations, lags behind behavioural learning, or occurs in tasks without reward [Gupta et al., 2010, Gillespie et al., 2021, Wittkuhn et al., 2025], suggesting a map-building function that stitches together separate experiences into a coherent relational model [Ólafsdóttir et al., 2018].

A critical limitation of prior work is that structure learning and value updating have rarely been studied simultaneously. In Liu et al. [2021b], for instance, participants were pre-trained on the task structure before the critical MEG session, so replay could be studied during value-based generalisation but not during the initial discovery of structure. Conversely, Wittkuhn et al. [2025] studied replay during structure learning but in a paradigm without reward, precluding investigation of value-based replay. This leaves open a fundamental question: when an agent must both discover latent structure *and* generalise value across it, which of these computational demands drives replay?

The empirical work presented in Chapter 5 addresses this question directly. Using fMRI and computational modelling, Renz et al. [2026] investigated replay during both the acquisition and adaptation of latent task structure in a value-based decision-making task. Participants learned that multiple sequential bandits shared a common latent reward source, a structure they had to discover from experience. Reward contingencies reversed periodically, creating moments requiring generalisation of newly observed outcomes to unobserved but structurally related options. On a second day, the latent reward structure itself changed, requiring flexible reorganisation of generalisation patterns. The central empirical questions were: (1) does replay occur selectively for task states sharing a latent reward source, even when those states have not been directly experienced, consistent with results by Liu et al. [2021b]? (2) does the strength and directionality of replay covary with structure learning and value generalisation? And (3) at the trial-by-trial level, are fluctuations in replay better explained by non-local value updates or by changes in inferred latent-cause assignments? These questions were addressed by combining the fMRI-based sequential slope ordering analysis [Schuck and Niv, 2019, Wittkuhn and Schuck, 2021] with a latent-cause inference model fitted to participants' behaviour [Gershman et al., 2010].

Chapter 3

Replay, Consolidation, and Representational Change During Sleep

The preceding introductory chapters established the theoretical and empirical foundations for understanding how neural replay supports learning and flexible behaviour during wakefulness. Chapter 1 introduced the complementary learning systems framework and the computational problem that replay helps to solve: enabling rapid hippocampal encoding without catastrophic interference in slower cortical networks. Chapter 2 reviewed the evidence for replay during wakefulness and its role in reinforcement learning, value propagation, and the construction of cognitive maps.

This chapter turns to the role of replay during sleep. While wakeful replay supports ongoing decision-making and the exploitation of learned structure, sleep replay has been proposed to serve a complementary function: the gradual redistribution of memories from hippocampus to neocortex, and the transformation of their representational content in the process. This idea, that sleep does not merely protect memories but actively reorganises them, is formalised in the *systems consolidation hypothesis* [Born et al., 2006, Diekelmann and Born, 2010, Klinzing et al., 2019], which provides the theoretical backbone of the empirical work presented. The present chapter reviews the origins of this hypothesis, the evidence for replay as its mechanistic implementation, and the emerging understanding of how consolidation changes what memories represent, not just where they are stored.

3.1 Theoretical Background

3.1.1 Systems Consolidation Hypothesis

The observation that recent and remote memories depend on different brain structures dates back to the clinical case of patient H.M., whose bilateral medial temporal lobe resection produced severe anterograde amnesia alongside a temporal gradient of retrograde memory loss: recent memories were severely impaired, while childhood memories remained largely intact

[Scoville and Milner, 1957]. This pattern suggested that the hippocampus is critical for the initial encoding and short-term retention of declarative memories, but that over time, memories become independent of the hippocampus and are instead supported by neocortical networks. Marr [1971] formalised this intuition in a computational framework, proposing that the hippocampus serves as a fast-learning temporary store that "teaches" the neocortex through repeated offline reinstatement of stored patterns.

The modern formulation of the systems consolidation hypothesis builds on the complementary learning systems (CLS) framework [McClelland et al., 1995]. CLS starts from the observation that effective learning faces a fundamental trade-off: a system that learns new information quickly risks catastrophically overwriting previously stored knowledge, a problem known as catastrophic interference. To resolve this, CLS proposes two complementary systems with different learning properties. The hippocampus functions as a fast-learning system that rapidly encodes specific experiences as sparse, pattern-separated representations, allowing individual episodes to be stored with minimal mutual interference. The neocortex, by contrast, is a slow-learning system that gradually extracts statistical regularities across many experiences, building up overlapping, distributed representations that capture the shared structure of the environment. The two systems are complementary precisely because their different learning rates serve different functions: fast hippocampal encoding preserves the details of individual experiences, while slow neocortical learning discovers what those experiences have in common.

The systems consolidation hypothesis extends this division of labour across time: memories initially depend on the hippocampus for storage and retrieval, but are gradually transferred to neocortical networks through a process of repeated offline reactivation. This transfer is thought to occur preferentially during sleep, when the absence of new sensory input creates a privileged window for hippocampal memory traces to be "replayed" to the neocortex without interference from ongoing experience [Born et al., 2006, Diekelmann and Born, 2010, Rasch and Born, 2013].

A critical prediction of this framework is that consolidation is not a passive copying process. Because hippocampal and neocortical networks have fundamentally different learning rules and representational formats, the act of transferring information between them should transform the memory's content [McClelland et al., 1995, Winocur et al., 2010]. Detailed, context-rich episodic traces in the hippocampus should gradually give way to more abstract, semantically organised representations in cortical networks. The seeds of this idea are already present in the original CLS framework: because the neocortex learns slowly by extracting regularities, information integrated into neocortical networks will naturally lose episodic detail and take on a more schematic character. The systems consolidation literature has made this prediction more explicit and empirically testable, emphasising that the representational format of a memory should change as a function of when and how it is retrieved [Winocur et al., 2010]. This prediction distinguishes systems consolidation from simple redundancy accounts in which hippocampal and cortical traces coexist independently, and it motivates the search for

representational changes, not just anatomical shifts, across consolidation.

Empirical support for the anatomical predictions of systems consolidation has come from multiple methodological traditions. In rodents, lesion studies established the classic double dissociation: hippocampal damage impairs recent contextual fear memories while sparing remote ones, whereas prefrontal and anterior cingulate lesions show the reverse pattern [Kim and Fanselow, 1992, Frankland et al., 2004, Maviel et al., 2004]. In humans, neuroimaging studies have traced the time course of this reorganisation. Gais et al. [2007] showed that when participants slept after learning word pairs, hippocampal activation during recall was enhanced at 48 hours, alongside increased functional connectivity between hippocampus and medial prefrontal cortex (mPFC). Six months later, memories encoded before sleep showed stronger mPFC activation compared to those encoded before sleep deprivation, indicating lasting changes in the neural architecture supporting retrieval. Critically, sleep deprivation on the first post-encoding night permanently disrupted this consolidation trajectory, demonstrating that the first night of sleep is essential for initiating systems-level reorganisation. Takashima et al. [2009] reported complementary findings: within 24 hours of learning face–location associations, retrieval-related hippocampal activity decreased while ventromedial prefrontal activity increased, with a corresponding shift from hippocampal–cortical to cortico–cortical functional connectivity. More recent work has demonstrated that this reorganisation can occur faster than classical accounts assumed: Brodt et al. [2018] showed that even a single night of sleep can establish a neocortical memory engram in the posterior parietal cortex, detectable through increased structural grey matter density and retrieval-related activation. At the cellular level, Ko et al. [2025] recently demonstrated that systems consolidation reorganises hippocampal engram circuitry in mice, providing convergent mechanistic evidence that consolidation physically restructures the hippocampal memory trace. Most recently, Ding et al. [2025] provided the first single-neuron evidence for overnight systems consolidation in humans: using intracranial microelectrode recordings and transformer-based neural network decoding, they showed that medial temporal lobe neurons could decode specific episodic memories before sleep but not after, while frontal cortex neurons showed the opposite pattern, a double dissociation that directly demonstrates the redistribution of memory representations from MTL to frontal cortex over a single night of sleep.

It is worth noting that the standard model of systems consolidation has been challenged and refined over the decades. The *multiple trace theory* [Nadel and Moscovitch, 1997] and its successor, the *trace transformation theory* [Winocur et al., 2010], argue that the hippocampus remains involved in the retrieval of vivid, context-rich episodic memories even at remote time points, and that consolidation produces a qualitative transformation from detailed episodic to gist-like schematic representations rather than a simple transfer. The *schema assimilation* account [Tse et al., 2007] further proposes that the rate and route of consolidation depend on the degree to which new information is compatible with pre-existing cortical knowledge structures (schemas): schema-consistent information may be rapidly integrated into neocortical net-

works, potentially bypassing prolonged hippocampal dependence. A further variant challenges the standard CLS assumption that fast and slow learning map cleanly onto hippocampus and neocortex, respectively. Schapiro et al. [2017] proposed that complementary learning systems may exist *within* the hippocampus itself: the trisynaptic pathway (entorhinal cortex $\rightarrow$ dentate gyrus $\rightarrow$ CA3 $\rightarrow$ CA1), with its sparse connectivity and high learning rate, supports rapid encoding of individual episodes, while the monosynaptic pathway (entorhinal cortex $\rightarrow$ CA1), with its more overlapping representations and slower learning rate, can extract statistical regularities across experiences. If the hippocampus already contains both fast and slow learning components, the division of labour during consolidation may be more distributed than classical accounts assume, with implications for where and how representational transformation occurs. Recent work provides complementary behavioural evidence that sleep can actively transform task representations: Löwe et al. [2025] showed that N2 sleep, but not N1 sleep, increased the likelihood of spontaneous insight in a perceptual task after a daytime nap. Exploratory analyses linked this effect to aperiodic EEG activity, consistent with theoretical accounts proposing that homeostatic synaptic downscaling during deeper sleep stages acts as a form of regularisation that suppresses irrelevant features and facilitates the discovery of hidden task structure.

These refinements and alternative models share the core prediction that consolidation transforms representational content, and they collectively motivate the empirical question of *what* changes about memory representations during sleep, not just where they are stored.

3.1.2 Replay During Sleep

The mechanistic account most commonly invoked to explain how memories are transferred from hippocampus to neocortex during sleep is neural replay. As reviewed in Chapter 2, hippocampal replay events recapitulate waking experience in temporally compressed sequences during sharp-wave ripples (SWRs; see Chapter 2 for a detailed treatment).

Chapter 2 introduced the *active systems consolidation* model as one of the major computational functions of replay [Born et al., 2006, Diekelmann and Born, 2010, Klinzing et al., 2019]. Central to this model is a coordinated dialogue between three cardinal oscillations of non-REM sleep: neocortical slow oscillations ($<$1 Hz), thalamocortical sleep spindles (12–16 Hz), and hippocampal sharp-wave ripples (80–120 Hz). During the depolarising "up state" of the slow oscillation, hippocampal memory traces are reactivated in the form of replay events that are temporally coupled to sleep spindles. The spindle, in turn, provides a temporal window during which arriving hippocampal information can be integrated into neocortical networks through spike-timing-dependent plasticity mechanisms [Klinzing et al., 2019]. This hierarchical nesting of oscillations, with slow oscillations driving spindles and spindles grouping ripples, is thought to provide the temporal scaffolding that coordinates hippocampal-to-cortical information transfer.

Substantial evidence supports this coordinated oscillatory framework. In rodents, hip-

pocampal replay during SWRs is temporally coupled to neocortical slow oscillations [Sirota et al., 2003, Isomura et al., 2006], and coordinated hippocampal–cortical replay during sleep reflects specific memory content rather than spontaneous co-activation, as demonstrated by the experience-dependence of inter-regional replay coupling [Ji and Wilson, 2007]. Recent work has further elaborated the circuit mechanisms: Karaba et al. [2024] identified a hippocampal circuit that actively balances the strength and frequency of memory reactivation during sleep, preventing runaway excitation, while Swanson et al. [2025] showed that the coupling between hippocampal SWRs and neocortical slow oscillations follows a topographic organisation consistent with directed, region-specific communication. Brodt et al. [2023] further emphasise that this unique electrophysiological milieu of non-REM sleep, characterised by reduced cholinergic tone, triple-nested oscillatory coupling, and altered excitation-inhibition balance, creates conditions that are more effective for systems consolidation than comparable reactivation during wakefulness.

An important temporal dynamic of sleep-dependent reactivation deserves mention. In rodents, replay rates are highest within the first 10–30 minutes of post-encoding sleep and then rapidly decline, a pattern consistent with the observation that the underlying neuronal ensembles are synaptically connected and that this connectivity is gradually weakened by the reactivation process itself [Brodt et al., 2023]. This temporal profile aligns with findings in humans showing that memory reactivations occur preferentially within the first 3–6 hours of sleep [Brodt et al., 2023], and it suggests that the initial period of post-learning sleep may be disproportionately important for initiating consolidation.

In humans, direct measurement of replay during sleep has been more challenging, but causal evidence for the role of sleep reactivation in consolidation has come from targeted memory reactivation (TMR) studies. In TMR paradigms, sensory cues that were associated with specific memories during encoding (e.g., odours or sounds) are re-presented during SWS, thereby biasing which memories are reactivated [Rasch et al., 2007]. Rasch et al. [2007] provided the seminal demonstration that odour cues presented during SWS selectively enhanced memory for items learned in the presence of that odour, an effect that was absent when cues were presented during REM sleep or wakefulness. Subsequent work has shown that TMR enhances hippocampal activation and hippocampal–cortical connectivity during sleep [Cairney et al., 2018] and can be combined with neurostimulation for more precise causal inference. Importantly, TMR does not act holistically on memory: Siefert et al. [2024] demonstrated that reactivation during sleep selectively enhances unique, distinguishing features of learned objects while weakening shared, category-typical features, consistent with a differentiation process rather than uniform strengthening. This finding provides causal evidence that sleep reactivation reshapes memory structure, not merely stabilises it.

Most strikingly, Geva-Sagiv et al. [2023] used closed-loop intracranial stimulation in epilepsy patients to synchronise prefrontal stimulation with the active phase of medial temporal lobe slow oscillations. This manipulation enhanced sleep spindles, improved coupling between hip-

pocampal ripples and thalamocortical oscillations, and boosted recognition memory, providing direct causal evidence that the precise temporal coordination of hippocampal–cortical activity during sleep supports memory consolidation.

An important distinction concerns what makes sleep replay particularly effective for consolidation, given that replay also occurs during wakefulness. As discussed in Chapter 2, wakeful and sleep replay appear to serve partially distinct functions, with wakeful replay emphasising online evaluation and updating of behavioural policies, and sleep replay emphasising systems-level consolidation. However, Brodt et al. [2023] argue that the critical distinguishing factor is not the replay phenomenon per se but the electrophysiological context in which it occurs: the triple-nested oscillatory coupling and neuromodulatory environment of non-REM sleep create conditions that preferentially drive hippocampal-to-neocortical transfer. The empirical chapter that follows examines one specific aspect of sleep's role: whether the neural signatures of sequential memory retrieval change across a night of sleep in ways consistent with the systems consolidation framework.

3.1.3 Overnight Representational Transformation

The preceding sections established that systems consolidation should transform the representational content of memories, not merely their anatomical locus. This section reviews the behavioural, neural, and computational evidence for such transformation.

Behavioural evidence for sleep-dependent abstraction

Sleep has been shown to promote the extraction of rules and regularities that are not explicitly taught during encoding. Wagner et al. [2004] demonstrated that participants who slept between encoding and test were more likely to discover a hidden shortcut in a number reduction task, suggesting that sleep facilitated the abstraction of an implicit rule. Ellenbogen et al. [2007] showed that sleep promoted relational memory integration, enabling participants to infer novel transitive relationships between items learned only in separate, overlapping pairs. Similarly, Durrant et al. [2011] demonstrated sleep-dependent consolidation of statistical learning, with sleeping participants showing stronger implicit knowledge of the underlying generative structure. However, the behavioural evidence is not entirely consistent, and critical review suggests that the conditions under which sleep promotes insight and generalisation are more constrained than initially assumed [Brodt et al., 2023]. The current consensus is that sleep can facilitate, but does not automatically guarantee, the extraction of higher-order structure, with the effect depending on encoding strength, structural complexity, and individual differences in sleep architecture. Recent work further suggests that sleep may preferentially rescue weak memories embedded in complex, interconnected networks, with the benefit scaling with fast sleep spindle density and amplitude [Schechtman, 2024].

Neural evidence for representational restructuring

At the neural level, multivariate pattern analysis has begun to reveal how memory representations are restructured across consolidation. Tompary and Davachi [2017] showed that neural patterns in both the hippocampus and mPFC came to represent the featural overlap across related memories, but only after one week of consolidation, not during immediate retrieval. Importantly, the degree to which individual memories preserved their unique encoding patterns was inversely related to the representation of overlap, pointing to a trade-off between episodic fidelity and integrative abstraction. Tompary and Davachi [2024] extended this finding to sequences with shared predictive structure: behavioural integration of overlapping sequences emerged only after a 24-hour delay and correlated with neural pattern changes in mPFC and with hippocampal–cortical resting-state connectivity.

Emerging evidence connects this representational restructuring to the successor representation framework introduced in Chapter 1. He et al. [2025] demonstrated that sequence learning reshapes neural representations to incorporate successor information, with representations of temporally adjacent stimuli becoming more similar even when temporal structure is task-irrelevant. Critically, both the strength of successor representations and their transformation toward higher-level abstract formats correlated with the proportion of slow-wave sleep during a post-learning nap, directly linking sleep architecture to the abstraction of relational structure.

Representational drift and constructive consolidation

A related phenomenon is *representational drift*: the gradual change in neural population codes for stable memories over time, even in the absence of new learning [Rule et al., 2019, Mau et al., 2020]. If neural codes are inherently unstable, then hippocampal-to-cortical transfer may involve re-encoding into a new representational format shaped by the neocortex's own computational constraints and prior knowledge, rather than the literal copying of a fixed trace. This constructive view of consolidation connects naturally to the empirical question addressed in the following chapter: whether retrieval-related neural patterns change in their regional distribution and representational content across a night of sleep.

Computational models of sleep-dependent consolidation

Building on the CLS framework (Chapter 1), several models have formalised how sleep consolidation transforms memory representations. Singh et al. [2022] developed a model of autonomous hippocampal–neocortical interactions during sleep, showing that NREM replay preferentially consolidates shared schematic features across memories, while the alternation of NREM and REM sleep stages enables continual learning by integrating new memories during NREM while protecting previously consolidated knowledge during REM. At the synaptic level, González et al. [2020] demonstrated in a biophysically realistic thalamocortical network

model that sleep replay orthogonalises overlapping memory representations via spike-timing-dependent plasticity, allowing the same neuronal populations to encode multiple distinct memories without catastrophic interference. Spens and Burgess [2024] proposed a complementary generative account in which consolidation corresponds to the training of a variational autoencoder through hippocampal replay. The generative network gradually learns the statistical structure of experienced events, producing schema-like representations that enable memory reconstruction and relational inference, but at the cost of systematic schema-based distortions that increase with consolidation. More recent models have incorporated schema-dependent consolidation rates [McClelland et al., 2020] and the role of replay in building predictive maps [Russek et al., 2017, Wittkuhn et al., 2025], converging on the prediction that replay-driven consolidation transforms the structure of stored memories.

Taken together, the evidence reviewed in this chapter points to a picture in which sleep, through the mechanism of coordinated hippocampal–cortical replay, actively transforms memory representations. The empirical chapter that follows tests specific predictions of this framework using a hierarchical sequence memory task with pre- and post-sleep fMRI retrieval sessions. The hierarchical task structure is designed to place differential demands on the two memory systems: fine-grained sequential recall requiring hippocampal pattern separation, and higher-order structural regularities amenable to neocortical statistical extraction. This design thereby provides a rich landscape for detecting consolidation-related changes in both the anatomical locus and the representational content of memory. As will be shown, sequential reactivation during retrieval is detectable across cortical and medial temporal regions, and its regional profile shifts across sleep in a manner consistent with the systems consolidation framework: parietal cortex emerges as the dominant locus of retrieval-related sequential reactivation after overnight consolidation, while replay strength predicts individual differences in retrieval accuracy.

Chapter 4

Measuring Brain Plasticity: Methodological Foundations

The preceding chapters have traced a progression from rapid, event-level representational updating to slower, overnight consolidation and transformation. Chapter 2 examined how neural replay during wakefulness supports the learning of latent task structure and propagates value across inferred states. Chapter 3 showed how replay during sleep drives the redistribution of memory representations from hippocampal to cortical networks, reshaping their content in the process. Both forms of replay-driven change, whether occurring within minutes during task performance or across hours during sleep, represent specific instances of a broader phenomenon: neural plasticity, the brain's capacity to reorganise its structure and function in response to experience.

This chapter steps back from those specific applications to survey the broader methodological landscape from which they emerge. The empirical chapters relied on a shared set of neuroimaging tools, in particular multivariate pattern analysis and representational similarity analysis, to detect and characterise representational change. But these tools are part of a much larger toolkit that neuroscientists have developed to measure plasticity across different spatial and temporal scales. Reviewing this broader landscape serves a dual purpose: it situates the analytic choices made in the preceding chapters within their wider methodological context, and it clarifies the constraints and trade-offs inherent in measuring neural change non-invasively in humans.

Plasticity, broadly construed, refers to the capacity of the nervous system to change its structure and function in response to experience, development, or injury [Paillard, 1976, Lövdén et al., 2010a, Pascual-Leone et al., 1991]. The concept has deep roots: Santiago Ramón y Cajal speculated that learning might involve the strengthening of connections between neurons [Rozo et al., 2024, Ramón y Cajal, 1894], and Donald Hebb later formalised this intuition in his influential proposal that coincident neural activity strengthens synaptic connections [Hebb, 1949]. Since then, plasticity has been demonstrated at virtually every level of neural organisation, from

molecular changes at individual synapses [Citri and Malenka, 2008, Malenka and Bear, 2004] to large-scale reorganisation of cortical maps [Maguire et al., 2006, Gärtner et al., 2013] and the growth of new neurons in the adult hippocampus [Boldrini et al., 2018, Eriksson et al., 1998, Disouky et al., 2026]. The replay-driven representational changes examined in the preceding chapters occupy an intermediate position within this hierarchy: they are too rapid to reflect gross structural reorganisation, yet they produce lasting changes in how information is encoded and retrieved, consistent with the functional consequences that define genuine plasticity.

In humans, direct measurement of synaptic or cellular plasticity is generally not possible. Instead, researchers rely on macroscopic, non-invasive methods that provide indirect but powerful windows onto structural and functional brain change. Structural neuroimaging techniques such as T1-weighted MRI and diffusion tensor imaging can reveal changes in grey matter volume or white matter integrity following training or intervention [Ashburner and Friston, 2000, Lövdén et al., 2010a, Sampaio-Baptista et al., 2018]. Functional methods, particularly fMRI, allow researchers to track how patterns of neural activity and functional connectivity reorganise as new knowledge is acquired [Poldrack, 2000, Bassett et al., 2011]. These approaches have produced landmark findings: London taxi drivers show enlarged posterior hippocampi related to years of navigational experience [Maguire et al., 2006]; musical training is associated with structural and functional differences in auditory and motor cortex [Gärtner et al., 2013, Sato et al., 2015, Schlaug et al., 1995]; and short-term learning paradigms can induce measurable changes in both grey and white matter within days or weeks [Sampaio-Baptista et al., 2018, Wenger et al., 2017].

Yet measuring plasticity in humans poses distinctive challenges. Observed changes can reflect genuine neural reorganisation but also vascular, metabolic, or methodological confounds [Poldrack, 2000]. Longitudinal designs are essential for tracking within-person change, but they introduce issues of test-retest reliability, attrition, and practice effects. Distinguishing plasticity from pre-existing individual differences requires careful control conditions and, ideally, random assignment to training versus control groups [Hariton and Locascio, 2018, Voss et al., 2010]. These challenges were not merely abstract concerns for the present thesis: the longitudinal within-person design of the sleep project, and the reliance on multivariate decoding rather than univariate contrasts in both empirical chapters, reflect deliberate responses to exactly these methodological demands.

This chapter is based on, and closely follows, the book chapter *Neuroscience Methods for Investigating Brain Plasticity* [Verra et al., 2025], published in The Oxford Handbook of Cognitive Enhancement and Brain Plasticity. In that chapter, my co-authors and I review structural and functional neuroimaging techniques, noninvasive brain stimulation approaches such as TMS and tDCS, and the study designs needed to draw valid inferences about plasticity. The chapter is reproduced here with minor formatting and referencing adjustments for consistency with the thesis.

Part III

Empirical Contributions

Chapter 5

Neural replay is connected to latent cause inference and supports fast generalization

Fabian M. Renz, Shany Grossman, Nathaniel D. Daw, Peter Dayan, Christian F. Doeller, Nicolas W. Schuck

Preprint published 16 January 2026: DOI: https://doi.org/10.64898/2025.12.12.693963

5.1 Introduction

A hallmark of adaptive behavior is the ability to generalize from past experiences to guide decisions in novel situations. This capacity relies on representations that capture the underlying decision-relevant structure of the environment while filtering out details irrelevant for reward prediction [Shepard, 1987, Niv, 2019, Kaplan et al., 2017]. When environments are novel or changing, successful generalization therefore requires to first infer an environment's hidden task states, a process often formalized as latent cause inference (LCI), which leverages similarities across observed events to infer the shared, unobservable causes that generated them [Collins and Frank, 2013, Lloyd and Leslie, 2013, Gershman and Niv, 2015]. Having established that distinct observations arise from the same latent cause, one can begin to generalize information (e.g. value) from one observation to the other.

A central question is how the brain initially learns and subsequently leverages such latent structure for flexible generalization. Recent work suggests a critical role for neural replay, the sequential reactivation of neural states corresponding to past experiences [Foster, 2017, Wittkuhn et al., 2021, Schuck and Niv, 2019]. However, the specific computational function of replay in learning and leveraging hidden structure remains debated, with research pointing to two distinct roles. One line of work suggests replay serves to update value expectations or plan immediate actions. For instance, rodent studies have shown that replay sequences often depict future paths to remembered goals, predicting immediate behavioral choices [Pfeiffer and

Foster, 2013]. In this view, replay utilizes the inferred structure to propagate reward information to related states, performing "non-local" value updating [Mattar and Daw, 2018, Liu et al., 2021b]. In contrast, alternative accounts of replay argue that replay is primarily concerned with memory, and in particular building and maintaining the abstract representation of the environment itself [Ólafsdóttir et al., 2018]. This view is supported by findings that replay often represents past goals or non-local areas rather than the immediate future path, decoupling it from current choices [Gupta et al., 2010, Gillespie et al., 2021], as well as by evidence for structure replay in tasks that do not involve reward learning [Wittkuhn et al., 2025]. Notably, Gillespie et al. [2021] found that replay often "lags" behavioral learning, suggesting it may function to consolidate the structure of the task, stitching together separate experiences to form a coherent map rather than drive immediate planning [Ou et al., 2025, Bakermans et al., 2025]. Recently, these views have been framed as complementary processes, where replay might prioritize either map-building or planning depending on the uncertainty of future goals [Sagiv et al., 2025, Ólafsdóttir et al., 2018].

Although methodological advances progressively allow the detection of replay in humans [e.g., Wittkuhn and Schuck, 2021, Wittkuhn et al., 2025, Liu et al., 2021b, 2019], empirical evidence distinguishing between these computational roles remains scarce, particularly during the initial discovery of task structure. Recent magnetoencephalography (MEG) work has shown that replay of unobserved paths can support value generalization across shared reward structures [Liu et al., 2021b]. This finding aligns with the "non-local" value updating account [Mattar and Daw, 2018], suggesting that replay utilizes offline periods to propagate new reward information across learned structure. Yet, because prior work has largely relied on pre-learned associations, it remains unknown how replay interacts with the discovery or adaptation of that structure. A critical mechanistic question persists: Does replay actively drive the inference of latent causes or does it primarily exploit an emerging structure to propagate value?

To capture the dynamical interaction between structure learning and value generalization and to identify its potential neural bases, we combine behavioral, neuroimaging, and latent cause inference modeling to investigate how latent structure is learned and used in the service of generalization. On the first day of our experiment, participants learned a novel graph structure linking multiple visual sequences to shared or independent reward outcomes, which repeatedly reversed throughout time (similar to Liu et al. 2021b). On the second day, we changed the latent reward structure, requiring flexible reorganization for appropriate generalization. This design allowed us to track both the acquisition of structure and its adaptation to change. To capture the intertwined structure and reward learning processes, we fitted a latent cause inference model to participants' data and, following our previous work [Schuck and Niv, 2019, Wittkuhn and Schuck, 2021, Wittkuhn et al., 2025], measured replay using functional magnetic resonance imaging (fMRI). Results indicate that participants learned the structure of our task and achieved successful generalization in the second half of day 1. A latent cause model fitted human behavior better than standard reinforcement learning models, successfully capturing

structure acquisition and generalization processes. Neural analyses in visual cortex revealed that replay was tightly aligned with participants' knowledge of the latent structure: its strength increased as the reward relationships were learned on day 1 and adapted after the structural change, including a pronounced shift in directionality for the newly generalizable path. These dynamics indicate that replay is closely coupled with the inferred structure and supports value updating, consistent with our LCI model in which trial-wise non-local value updating of the unobserved path was the strongest predictor of replay. Moreover, we found that multivariate patterns in the medial temporal lobe reflected the emerging reward structure and reorganized following structural change.

5.2 Results

5.2.1 Structure learning facilitates fast value generalization

Fifty-two healthy participants (mean age 27.4; 25 females) completed a 4-session fMRI experiment across two consecutive days (Fig. 5.1D). The primary purpose of our study was to investigate the role of neural replay in (a) learning the latent task structure and (b) the structure-enabled value generalization.

In both days of the main task, participants had to decide between four sequential bandits (A, A*, B and C), each consisting of a sequence of four 750 ms videos that led to a reward outcome (e.g., A1 $\rightarrow$ A2 $\rightarrow$ A3 $\rightarrow$ A4 $\rightarrow$ R1; videos depicted distinct objects e.g., giraffe, grapes, basketball; see Fig. 5.1A and SI Fig. 5.11). Each sequence terminated in a scalar outcome (0–100 points), with rewards for three bandits (A, A*, and B) sampled from one latent reward distribution (R1) and rewards for bandit C sampled from a second distribution (R2); these distributions were noisily sampled on each trial and periodically reversed such that which latent reward was currently higher changed over time. The video sequences of bandits A and A* depicted objects from the same categories (e.g., A1 and A*1 depicted a giraffe and a zebra, i.e. both animals; the other matched categories were houses, cars, and fruits), whereas the content of all other videos was unrelated. On each trial, participants first passively observed one of the four video sequences and its outcome, and were then asked to choose between two video screenshots using a preference slider (1-6). The choice options were always matched in their positional distance from the reward (e.g. B2 vs. C2, see Methods for details). Feedback on overall accumulated reward was provided every 10 trials. The task involved 120 trials structured into 5 blocks, and reward outcomes reversed episodically every 15-20 trials (i.e., 6 reversals in total, see Methods).

As expected, participants successfully learned reward contingencies and adjusted their choices upon reversals to the bandits with the highest outcomes (see Fig. 5.2A; remaining reward schedules in SI Fig. 5.8), demonstrated by an increase in accuracy from 68.7 % accuracy in Block 1 to 84.8 % in Block 5 (linear effect of block in a linear mixed effect model: $\beta = 0.041$, $z = 6.46$,

$p < .001$, see Fig. 5.2B).

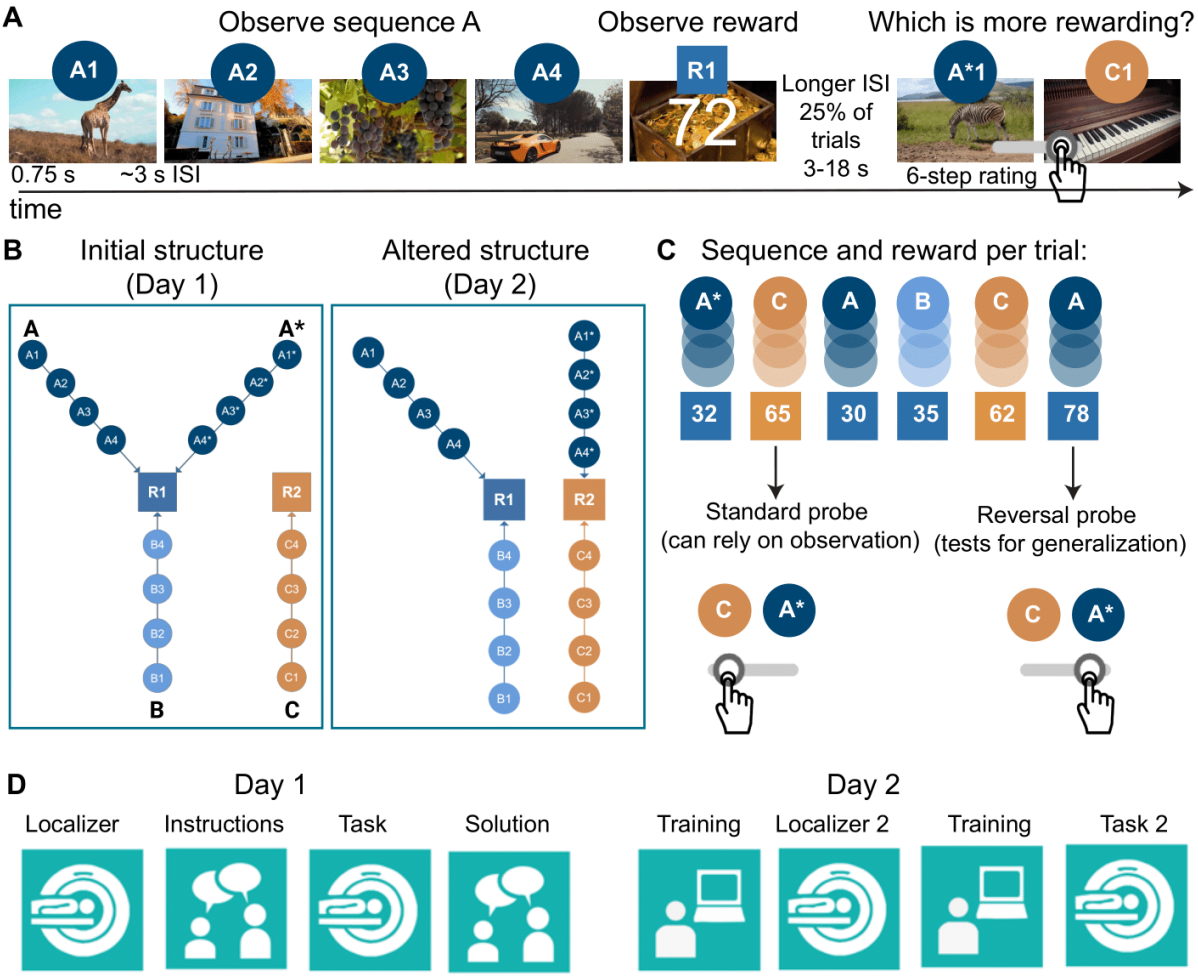

Figure 5.1: **Task design and experimental paradigm.** (A) Each trial presented a sequence of four 750 ms video stimuli leading to a reward outcome (points), and was then followed by a choice between two options using a 6-point confidence slider. (B) On Day 1, three sequences (A, A*, B) led to a shared reward outcome (R1) and sequence C to an independent outcome (R2); A and A* shared the semantic categories for each sequence position (e.g. animals: giraffe and zebra). On Day 2, after 40% of the trials, sequence A* covertly switched from R1 to R2, requiring participants to learn the new underlying structure and adapt their generalization strategy accordingly. (C) Trials were either *standard trials*, in which values could be judged from direct observation, or *reversal trials*, which tested generalization to unobserved sequences after a reward reversal. Correct structure knowledge enabled generalizing value changes from an observed sequence (e.g., A) to others sharing its reward contingency (e.g., A* and B). (D) Experimental timeline: Each day began with an fMRI localizer used for classifier training, followed by task training and the main task. Day 1 involved structure discovery from experience, ending with explicit revelation of the structure. Day 2 therefore began with full structure knowledge, allowing assessment of exploitation and adaptation to the mid-session structure change.

A core feature of our task on Day 1 was that, unbeknownst to participants, the outcomes of three sequential bandits (A, A* and B) were sampled from the same latent mean (shared reward, R1), while bandit C was associated with an independent reward (R2, unique reward, Fig. 5.1B). Our main behavioral question was whether participants would learn about the correlation structure of the rewards and use it to generalize outcome observations among all three related bandits. To test this, we used trials immediately after a reward reversal to provide a primary measure for participants' knowledge of the graph structure. Specifically, we asked

participants to choose between two nodes of the graph that had not been observed since the reversal. To answer correctly in these reversal-locked trials, participants had to generalize value from an observed bandit (e.g. A) to one of the correlated bandits (e.g., A*), based on shared reward contingencies rather than direct observation (henceforth reversal trials, see the right choice in Fig. 5.1C). Note that in these trials, responding based on one's memory of a bandit's last observed reward would lead to objective accuracy of less than 50%: Although the last rewards observed for bandits A* and C indicated that C is better than A*, for instance, in reversal trials participants should prefer A* over C because they had just observed that A resulted in a higher outcome on the last trial.

An analysis of reversal trials on Day 1 showed that participants improved accuracy from around 55% in blocks 1-2 to 74% in blocks 4-5 (Fig. 5.2C, main effect of block in lme: $\beta = 0.052$, $z = 3.041$, $p = 0.002$; performance in blocks 4 and 5 was above chance, both p < 0.05, FDR corrected). This pattern was also evident in the trial-wise accuracy following reward reversals (Fig. 5.2D). During early reversals (first 2), accuracy showed a maximum drawdown to 41% immediately after the reversal and did not recover to reliably above-chance performance within the subsequent generalization trials (post-reversal trial 4: 60.8%, $p = 0.02$; trial 5: 58.0%, $p = 0.11$). In the middle phase of the task (reversals 3 and 4), the maximum drawdown decreased to 49%, following which participants returned to above-chance performance by the fifth reversal generalization trial (post-reversal trial 4: 62%, $p = 0.106$; trial 5: 68%, $p = 0.009$). During the late stage of the task (reversals 5 and 6), the maximum drawdown was 51%, after which participants performed above chance from the second post-reversal trial onward (post-reversal trial 2: 67%, trial 5: 84%; all $p < 0.001$ from trial 2 onward, FDR-corrected; Fig. 5.2D). The slight dip before the reversal likely reflects participants beginning to adapt as rewards start changing but have not yet reversed, as reversals transition through 1–2 intermediate steps between stable reward levels (see Methods).

As mentioned above, videos associated with bandits A and A* were semantically similar, while videos in A/A* and B were not. We hypothesized that semantic similarity would facilitate generalization [e.g., Tompary and Thompson-Schill, 2021]. Indeed, reversal trials requiring generalization between category-sharing sequences (e.g., from A to A*) showed superior performance compared to cross-category generalization (e.g., from B to A*; paired t-test $t(50) = -2.56$, $p = 0.014$, see Fig. 5.2E). Hence, while generalization occurred for bandit pairs with and without semantic similarity, such similarity did enhance structure learning and generalization.

We also included 10 trials that probed two options that led to the same reward (e.g., A2 vs. A*2, both leading to R1), which therefore had no "correct" solution under the initial structure. On these trials, participants showed significantly weaker preferences than on trials involving options from different reward sources, which elicited more extreme ratings. (A linear mixed-effects model predicting rating extremity (center = low, extreme = high; trial type $\beta = 0.343$, $z = 13.26$, $p < .001$).

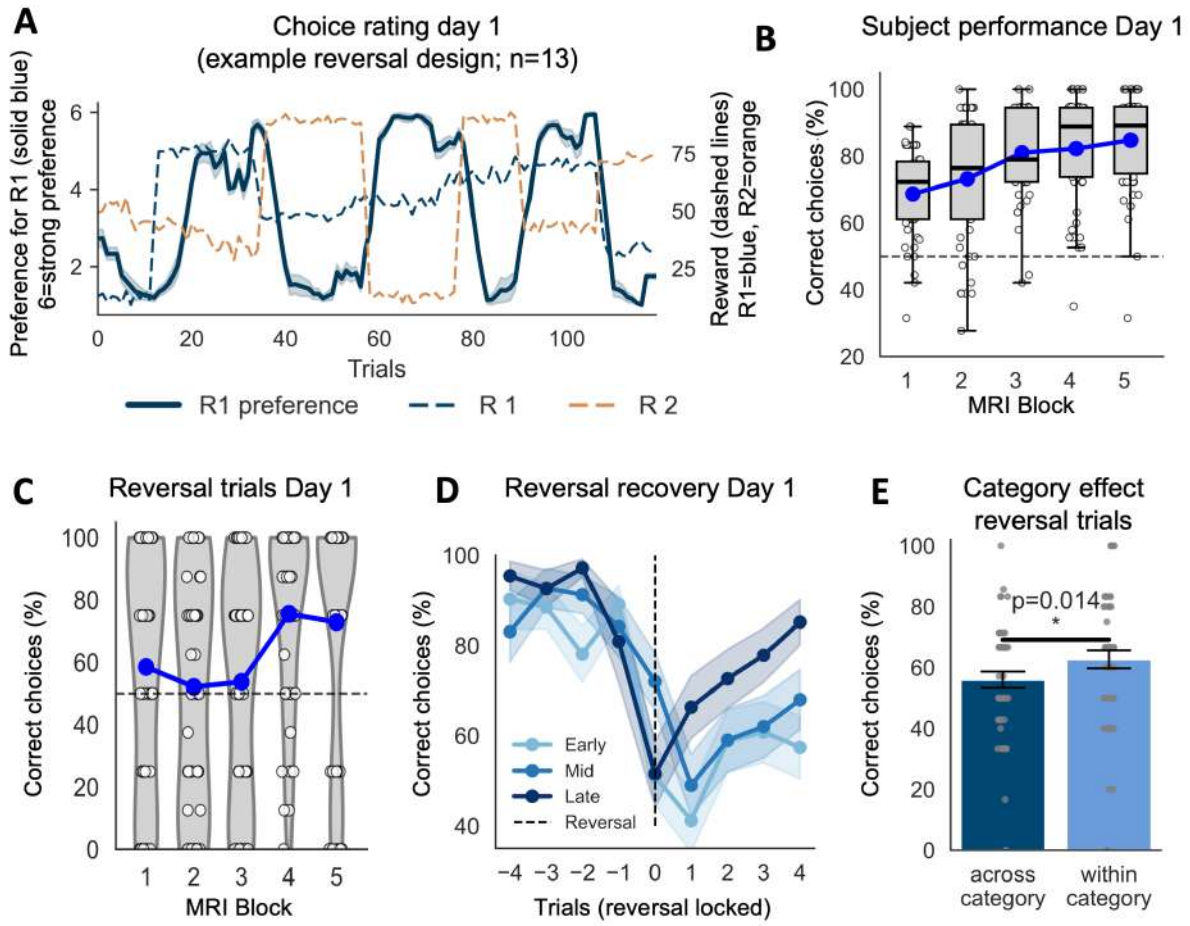

Figure 5.2: **Behavioral performance on Day 1 demonstrates structure learning and adaptation.** (A) Example preference ratings for one reward schedule (solid blue line; 1–6 scale) for R1 compared with the true reward values (dashed lines; R1=blue, R2=orange) across Day 1. Participants successfully adapted to the reward reversals: when R1 $>$ R2, ratings for R1 increased, indicating a strong preference, whereas when R2 $>$ R1, ratings for R1 decreased, reflecting a shift in preference toward R2. (B) Accuracy across all trials, showing progressive improvement over the course of Day 1. (C) Performance on reversal trials that required generalization to unobserved sequences. Accuracy begins at chance early in training and increases with experience, reaching above-chance levels in blocks 4 and 5. (D) Recovery from reversals, time-locked to the reversal onset (trial 0), shown for early (reversals 1 and 2), mid (reversals 3,4), and late reversals (reversals 5,6) across Day 1. Recovery becomes faster over the course of training, indicating improved use of latent structure. (E) Semantic similarity effect on Day 1 reversal trials: performance was higher when generalizing within category-sharing sequences (e.g., observing A and being probed on A*).

5.2.2 A latent cause inference model captures structure discovery and generalization

One mechanism that can explain structure learning and value generalization is a process known as latent cause inference (LCI). This framework posits that the brain infers hidden, unobservable causes to explain the structure of observed patterns of events and then generalizes on the basis of inferred latent causes [Gershman et al., 2010, Lloyd and Leslie, 2013, Gershman and Niv, 2015]. To test this idea, we developed a LCI model of our task that used a Chinese Restaurant Process prior with particle filtering to perform approximate inference [see Sanborn

et al., 2010, see Methods] about hidden causes that might have generated the observed video sequences and their rewards. Different particles represent alternative hypotheses about latent cause structures and are resampled according to how well they account for the observed data. Because the model continuously inferred the structure of latent causes given past observations, it could learn about reward correlations and generalize values after reward changes by discovering shared latent causes across paths (Fig. 5.3A).

We applied the LCI model (4 free parameters: concentration parameter α, categorical update η, hazard rate h, and choice temperature θ) to data from our task and tested it against three alternative models: a simple temporal difference (TD) model that learned about each bandit independently (two free parameters learning rate α, temperature θ), and two versions of 'clairvoyant' TD models that applied each TD update to all bandits which shared outcomes with either a single or differential learning rates (the models are clairvoyant because they magically know from the start how outcomes can be generalized). The TD model had the worst model fit (AICc = 69.6), due to its failure to generalize following reversal trials (see SI). The clairvoyant TD models had marginally better model fits because they did generalize (single learning rate AICc = 63; differential learning rates AICc = 56.6), but their model fit was limited by their inability to capture the learning progression throughout the first day (as well as the structure change during the second day, see below). The LCI model's flexibility in both discovering and updating structure yielded the best fit (AICc = 43.8; Fig. 5.3D). Inspecting the LCI model more closely confirmed that it captured key behavioral phenomena across both days (e.g., reversal trial generalization, see Fig. 5.3B; reversal recovery SI Fig. 5.9C).

The latent cause model also captured individual differences in structure learning. Participants with higher overall accuracy were fit by models discovering fewer latent causes (Fig. 5.3C; linear mixed-effects model: $\beta = -0.022$, SE $= 0.006$, $z = -4.03$, $p < 0.001$), indicating efficient structure discovery (i.e., value generalization was achieved by grouping sequences with shared rewards under a common latent cause). Moreover, the selectivity of value updating within the model correlated with participants' reversal performance: participants' accuracy, calculated as the percentage of correct choices in reversal trials, was positively associated with the degree to which the model exhibited higher non-local updating during reversal compared to stable periods. (LME: $\beta = 0.021$, $z = 2.825$, $p = 0.005$; Fig. 5.3D). Together, these results suggest that statistical learning supports the inference of latent structure that scaffolds value generalization, enabling selective, non-local value updating when contingencies change while preserving stability during steady states.

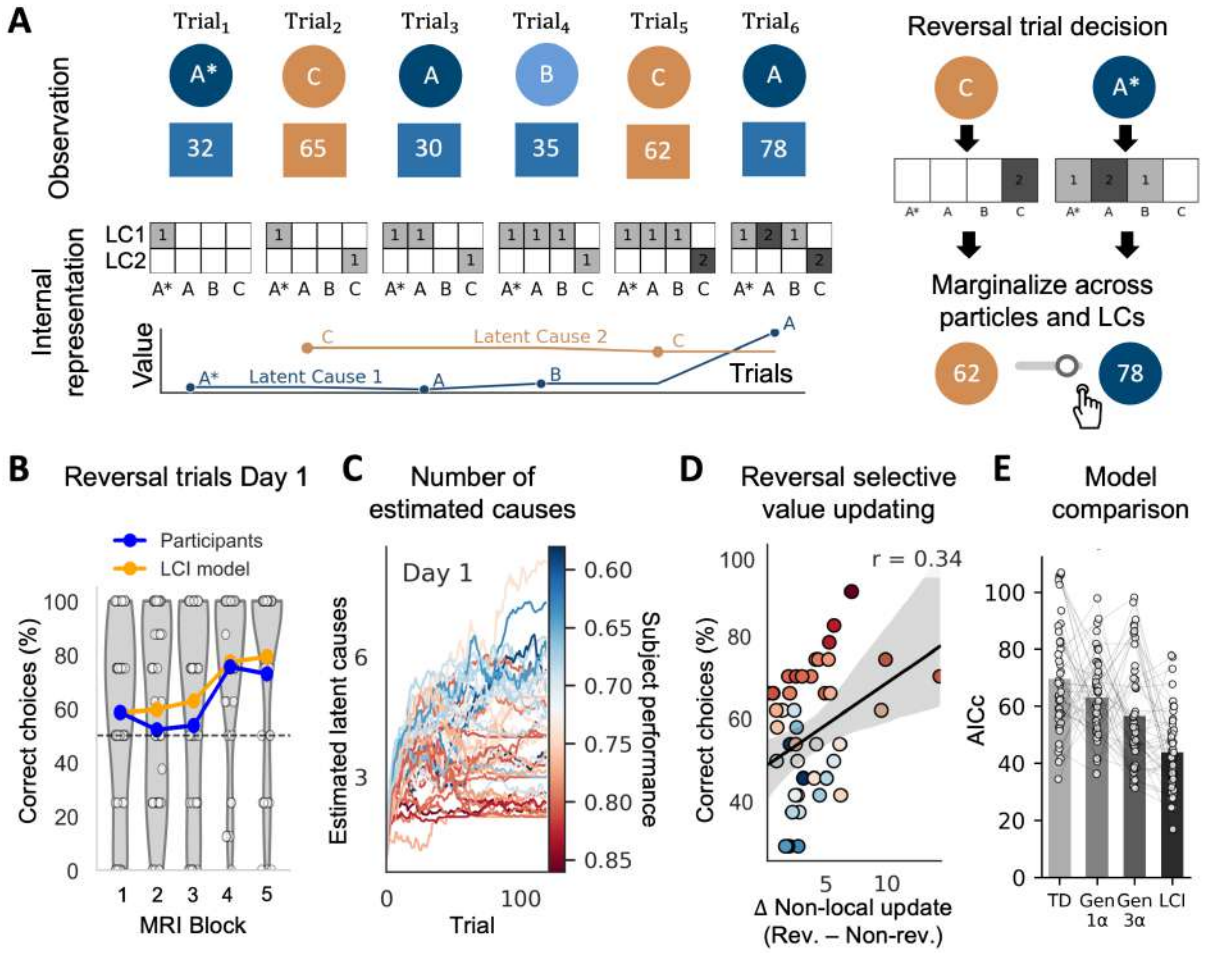

Figure 5.3: **Latent cause inference model explains structure learning on Day 1.** (A) Illustration of the latent cause inference model. On each trial, the model computes the likelihood of the current observation for all existing versus a potential new latent cause, combining (i) how often the observed sequence has previously been assigned to that cause and (ii) the cause's current reward expectation. The likelihood is combined with a Chinese Restaurant Process prior, which favors reuse of frequently inferred causes. Particles then branch over all possible assignments and are resampled in proportion to how well each branch accounts for the observed data. As sequences become associated with shared latent causes, updated values assigned to one sequence can be accessed by other sequences through their shared latent-cause representation, enabling generalization. Trial-by-trial value estimates are obtained by marginalizing over latent causes and particles. (B) Reversal trial performance calculated as percent correct choices on Day 1. The LCI model captures the gradual improvement in participants' generalization accuracy across blocks. (C) Individual differences on Day 1. Participants with higher behavioral performance were best fit by models inferring fewer latent causes, consistent with more efficient structure discovery and generalization through shared latent causes. (D) Individual differences in selective value updating relate to reversal performance. The x-axis shows the model-derived increase in non-local value updating for unobserved sequences during reversal trials relative to non-reversal trials (Rev. − Non-rev.). The y-axis shows participants' reversal performance, calculated as the percentage of correct responses. Greater model-derived selective value updating was associated with higher behavioral performance. Point color reflects overall task accuracy, with warmer colors indicating higher performance (same color scale as panel C). (E) Model comparison across the full experiment. The LCI model outperforms both temporal-difference models lacking generalization and models with fixed, non-adaptive generalization structure. This performance difference is preserved when the analysis is conducted separately for each day (see SI Fig. 5.9A).

5.2.3 Backward replay selectively reactivates unobserved generalizable paths

Following previous work [Liu et al., 2021b], our core neural hypothesis was that value generalization was facilitated by neural replay, i.e. that generalization from one bandit to another was driven by post-outcome backward replay of the bandit to which outcome knowledge was generalized. In line with our behavioral findings, we specifically expected such "non-local" replay of unobserved sequences to occur for bandits that shared reward contingencies. To test this hypothesis, we inserted 18-second breaks after reward presentation, just before the decision phase, i.e. the moment when value generalization would be needed to support subsequent choices. These prolonged inter-stimulus intervals (ISIs) were distributed throughout the task but had systematically higher frequency after trials following reward reversals, as these were the critical moments during which non-local updating was most needed for successful decision-making (40 per session). Almost all reversal trials were associated with prolonged ISIs and we did not observe any performance difference comparing non-reversal trials with and without prolonged ISI ($t(50) = 1.393$, $p = .170$).

To decode replay-related neural signals, we first trained logistic regression classifiers on fMRI patterns from independent localizer sessions conducted before and after the task (sessions 1 and 3, see Methods and SI Fig. 5.10). Classifiers were trained in a one-versus-rest fashion to distinguish between the 16 individual video stimuli. We first focused on a visual region of interest (ROI; object-selective regions in the ventral stream, as in Wittkuhn and Schuck [2021], for details see Methods), which yielded a low between-class confusion (see Fig. 5.4A) and very high classification performance (mean probability of true class at peak: 71.2%, $t(49) = 44.02$, $p < .001$, Fig. 5.4B) that is a prerequisite for subsequent replay analyses. Note, one classifier was excluded from all replay analyses because it showed elevated activation during the reward stimulus (see Methods), which could have biased the prolonged ISI. Importantly, the exclusion of this classifier did not selectively impact any condition. Because condition assignments depended on the presented sequence, the sequence containing the excluded classifier was represented in all four conditions introduced below (see Fig. 5.4D).

Classifiers trained on localizer sessions also successfully decoded the displayed stimuli during the main task (probability at peak in the visual ROI: 51.1%, $t(50) = 31.35$, $p < .001$, Fig. 5.4C and SI Fig. 5.12). To study replay during the ISI periods, we employed the sequential slope ordering analysis (SODA) method developed and used in Schuck and Niv [2019], Wittkuhn and Schuck [2021], Wittkuhn et al. [2025]. The main idea of SODA is that sequential neural activity will lead to a specific pattern of ordered classifier probabilities across time, which can be detected by analyzing the slopes of TR-wise regression coefficients that test a hypothesized sequential order. Hence, we could test for instance, whether class probabilities of classifiers of the A branch were ordered as expected ($p(\mathrm{A1}) > p(\mathrm{A2}) > p(\mathrm{A3}) > p(\mathrm{A4})$), and similarly for the other three sequences. The slope assesses average ordering, and thus does

not guarantee that every comparison is conjunctively significant. We excluded the reward state from the sequence modeling, as it was shared across all four sequences. Given the set of possible within-branch orderings, we constructed for each ISI period four sequence slope models and categorized them according to the branch that had been observed in the trial immediately before that ISI (Fig. 5.4D). First, sensory model coefficients referred to the branch just seen before the ISI and hence reflected sequential activation potentially arising due to the actually observed video sequence. Second, the semantic generalization model quantified sequential reactivation of the bandit that is semantically related to, and shares a latent reward with, the bandit shown before the ISI (e.g., reactivating the sequence A4→A3→A2→A1 after observing A*), if any; Third, non-semantic generalization model slopes reflected reactivation of sequences that share the same latent reward, but not the semantic category (e.g., replaying the sequence B4→B3→... after observing A*). Finally, the non-generalization model captured reactivation of sequences that are unrelated to the just observed one (e.g., C4→C3→... after observing A*). We hypothesized that we would observe Semantic and Non-Semantic generalization effects that indicate backward replay. Sensory forward activation elicited by the presented video sequence served as a sanity check for our analyses, and Non-generalization sequences functioned as a control condition.

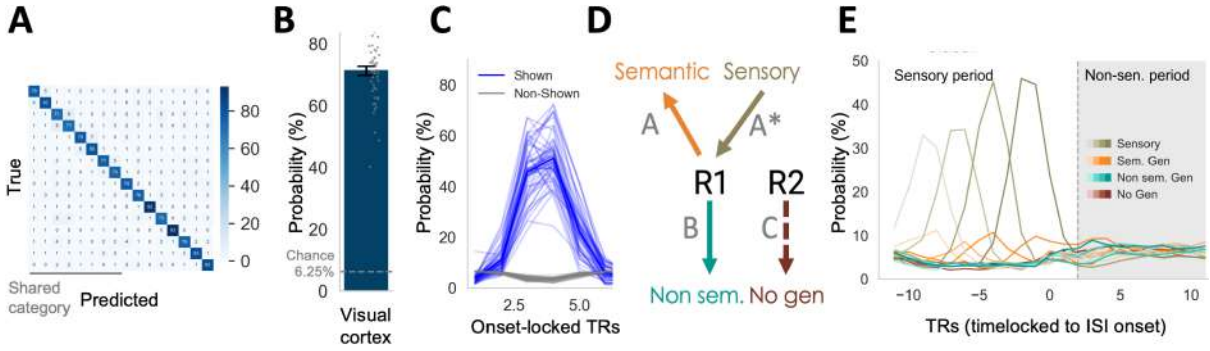

Figure 5.4: **Training and applying classifiers** (A) Confusion matrix from localizer sessions showing accurate decoding of 16 video stimuli in visual cortex with minimal category bias. (B) Cross-validated localizer performance in visual cortex (avg. probability of true class 71.2%; chance level 6.25%, dashed line). (C) Transfer to the main task: classifier probabilities peak following stimulus onset for shown items (blue) but not non-shown items (gray). (D) Replay conditions. After viewing sequence A* (sensory, olive), unobserved sequences were classified as semantic generalization (A, shared reward and category, orange), non-semantic generalization (B, shared reward only, turquoise), or no generalization (C, unrelated, brown). (E) Averaged classifier probability time courses aligned to prolonged ISI onset (time 0). Strong sequential sensory responses precede ISI onset, followed by sequential reactivation during the non-sensory period. The sensory period refers to the initial TRs dominated by stimulus-driven responses, whereas the non-sensory period begins once these sensory responses return to baseline.

Focusing on the visual ROI, we found strong evidence for sensory-driven sequential activation during the presentation of the sequence and the early TRs of the prolonged ISI, as expected (Fig. 5.4E). As shown in Fig. 5.5A, the regression slope testing the sensory sequence showed a sinusoidal pattern, with positive coefficients (indicating forward ordering, βs > 0 at TRs -11–6 relative to the ISI onset, all ps < 0.01, corrected) followed by negative coefficients (reflecting

the expected sign flip that is part of any forward sequence, see [Wittkuhn and Schuck, 2021], βs < 0 at TRs -4–1, all ps < 0.01, corrected). Hence, our classifiers were able to detect objective sequential activation.

We also observed that unobserved but semantically related categories co-activated during the stimulus presentation period, although confusion among these semantically related categories during the localizer was minimal (Fig. 5.5A). To test whether this effect indeed reflected meaningful co-activation or merely classifier uncertainty, we quantified how the classifier distributed residual probability across non-presented classes during the localizer versus main task. For each trial, we computed the difference between (i) the probability assigned to the unobserved but semantically-related class and (ii) the mean probability assigned to all other unobserved classes that were not part of shown or related category sequences. If co-activation resulted solely from classifier uncertainty, the magnitude of this difference should remain the same in the localizer and main task, while associative links between category members as reported previously [e.g., Wimmer and Shohamy, 2012] would lead to a larger difference in the main task. A linear mixed-effects model confirmed a significant increase in category-partner bias during the main task relative to the localizer ($\beta = 0.009$, $z = 2.65$, $p = .008$), indicating that unobserved but semantically related items were activated more strongly than can be accounted for by the baseline classifier confusion.

We next asked whether we could detect replay in the visual ROI during the post-sensory period. We conservatively defined this period as the TRs when the coefficients of the above-described sensory model had returned to 0, which was true starting from 2 TRs into the prolonged ISI (mean slope: 0.0005; t-test against 0: $t(50) = 0.233$, $p = .81$), see vertical line in Fig. 5.5A,B). To test whether generalization-related visual replay occurred in this period, we examined the β coefficients of *condition* i.e., the semantic and non-semantic generalization models, in comparison to the no generalization and sensory models. A linear mixed effects model of average β slopes across the entire post-sensory ISI period with condition as a predictor revealed significant differences across sequence conditions (main effect condition: $z = 5.89$, $p < 0.001$, see Fig. 5.5C). Post hoc tests showed that both semantic generalization ($t(49) = -2.55$, $p = 0.028$) and non-semantic generalization ($t(49) = -3.27$, $p = 0.008$) exhibited significant negative effects, indicating systematic backward replay. In contrast, no evidence was found for sensory replay ($t(49) = 0.76$, $p = 0.600$) or for no-generalization sequences ($t(49) = 0.16$, $p = 0.876$; Fig. 5.5C). Direct comparisons further confirmed that sensory replay was significantly weaker than both non-semantic generalization ($t(50) = 4.10$, $p < 0.01$) and semantic generalization ($t(50) = 3.81$, $p < 0.01$). We found no evidence for differences between semantic and non-semantic generalization, either in the overall post-sensory effect ($t(50) = -0.26$, $p = .80$) or at individual TRs. Within the non-sensory window, the smallest uncorrected p-value occurred at TR 5 ($t(50) = 1.98$, $p = .053$). After correcting for multiple comparisons, however, no time point reached significance (all $p > .24$).

Having established that visual replay occurs selectively for generalizable rewards, we next

asked whether trial-by-trial fluctuations in replay on Day 1 predicted subsequent choices and generalization. Averaging behavioral accuracy across the following five trials, replay strength predicted better performance ($\beta = -0.036$, $z = -2.38$, $p = .017$). Replay strength did not predict accuracy on the immediately following decision, however (mixed-effects model: $\beta = 0.029$, $z = 1.03$, $p = .302$). Given that participants first had to learn about the reward structure before they could begin to generalize, we also tested whether generalization replay increased in strength over time. This was indeed the case (main effect of block $\beta = -0.001$, $z = -2.7$, $p = 0.007$, across semantic and non-semantic generalization), supporting the idea that the detected visual replay reflects the learned cluster structure. We did not observe any difference between semantic and non-semantic generalization in the increase over time (interaction of block and condition in lme: $\beta = 0.002$, $z = 1.59$, $p = 0.112$).

Stronger replay for generalizable (unobserved but reward-related) sequences compared to unrelated sequences, increase in selective replay over time, and a link to future performance demonstrate that replay in visual cortex was not driven by recent stimulus exposure, but instead reflected the reward learning based on participants' knowledge of the task structure.

5.2.4 Non-local replay is linked to value reversals

Having established the selective and dynamic nature of visual replay, we sought to determine which factors triggered it. Prior theoretical and empirical work has indicated that a main function of replay could be to facilitate value updating [Liu et al., 2021b, Schaul et al., 2015, Mattar and Daw, 2018, Sinclair and Barense, 2019], and emphasized the effect of prediction errors, which in our study occurred predominantly following reward reversals. Some work has emphasized the role of replay specifically in *non-local* value updating, which uses observed prediction errors to update non-observed states to which values can be generalized [Liu et al., 2021b]. By linking three video sequences to the same outcome on Day 1, our task encouraged fast value generalization, such that replay would be expected following reversals from a non-local value updating perspective.

Given that we found no evidence for replay in the non-generalization condition, we focused on modeling the average trial-wise slopes of the semantic and non-semantic replay models (averaged over all post-sensory TRs) as a function of period (5 trials following a reversal, vs. stationary periods, see Methods) and replay condition (semantic vs. non-semantic). Slopes were averaged within each trial across the TRs of the non-sensory period. This analysis showed indeed that reversal periods were a significant predictor of replay (main effect period: $z = -5.218$, $p < 0.001$; Fig. 5.5D). Model comparison to a baseline model without reversal period also showed significantly improved model fit for the full model with a reversal factor (likelihood-ratio test: $\chi^2(2) = 19.18$, $p < .001$).

5.2.5 Neural replay aligns with latent cause inference dynamics

While the observation of reversal-triggered replay is interesting, reversals are an experimenter-defined task aspect that participants can only infer indirectly. Hence, the link between task period and replay falls short of providing a link between a genuine cognitive variable and replay. We therefore used the latent cause models that were fitted to participant's behavior and extracted two cognitive variables of interest that could drive replay, with the goal of providing a statistical model of learning-related trial-by-trial fluctuations in replay. We focused on two main aspects of learning that have been proposed to play a role in replay: nonlocal learning to attribute value to different causes [Mattar and Daw, 2018], and learning of the latent cause structure itself [Sagiv et al., 2025]. To test these hypotheses, we used the model to estimate the magnitude of value and structure learning occurring at each trial, and asked whether these correlated with replay. While our model is agnostic to the actual process by which these aspects of learning are implemented in the brain, it served as a way to statistically disambiguate our alternative hypotheses.

We quantified the amount of non-local value change that occurred in each participant's model following each reward observation, i.e., how much the model updated the value estimate of A after observing B and so forth (Fig. 5.5E). Contrasting the non-local value updates in our model for reversal trials vs. stationary periods showed that in our model average non-local value updates were larger during reversal periods than during stable periods ($t(46) = 9.541, p < .001$), suggesting that this might be a plausible driver of replay. To quantify trial-wise structure learning, we examined how similar the model's latent-cause probability distributions were across sequences. For every sequence pair and each trial, we computed the Kullback–Leibler (KL) divergence between their latent-cause distributions. Lower divergence indicates that the model assigns similar latent-cause probabilities to both sequences, effectively mapping them onto the same underlying structure. As shown in Fig. 5.5F, sequences that shared a reward exhibited consistently lower divergence, whereas non-shared sequences showed higher divergence. To align this measure with our trial-resolved value-update analysis, we derived the trial-by-trial change in divergence, capturing how strongly the model updated its structural beliefs on each trial. This trial-wise update served as a proxy for ongoing structure learning, in direct parallel to the model's non-local value updating.

Having computed these two scores, we asked to what extent they could explain trial-by-trial replay fluctuations in the visual ROI (Fig. 5.5G). A linear mixed effects model that predicted fluctuations in replay of each non-observed path after each trial showed that including the non-local value change variable as a predictor improved model fit significantly over a baseline model with only a factor for replay condition (AICs: -8375.9 vs. -8373.6; likelihood ratio test value change variable and condition vs. condition only: $\chi^2(2) = 6.39, p = 0.041$). Note that the metric we included from the LCI model made predictions not only about fluctuations in overall replay strength, but included separate expectations for each arm, i.e. how strong

replay of A* vs. B states was in a given trial, depending on the individual learning history of a given participant up to the trial in question. Using a more generic model that predicted the same replay strength for each arm did not reach statistical significance (likelihood ratio test averaged value change variable and condition vs. condition only: $\chi^2(2) = 5.488, p = 0.064$). Including the model-derived measure of structure learning, in contrast, did not improve model fit over the baseline condition-only model either (AIC -8373.6, likelihood ratio test: $\chi^2(2) = 3.999, p = 0.135$). Thus, replay in our task seemed mainly related to non-local value fluctuations, which we could predict in a trial and path specific manner from our model. Note that non-local value fluctuations and reversals were correlated, and we did not find evidence that the non-local value model performed better than a model that incorporated only reversal period, i.e. adding non-local value change to a model that already included a regressor for task period did not yield a significant improvement (AICs: -8377.8 vs. -8379.3, likelihood ratio test: $\chi^2(2) = 2.433, p = 0.296$).

Lastly, we applied the same decoding and sequentiality analyses to classifier evidence from an anatomically defined medial temporal lobe ROI (MTL; hippocampus, parahippocampal cortex, and entorhinal cortex), motivated by the hippocampal system's established role in replay in rodents [Wilson and McNaughton, 1994, Skaggs and McNaughton, 1996, Ólafsdóttir et al., 2018] and MEG evidence which appears to indirectly localize human replay signals to hippocampal regions [e.g., Liu et al., 2019, 2021b]. Decoding in MTL was lower than in visual cortex, but remained reliably above chance (localizer peak true-class probability: 11.5%, chance: 6.25%; $t(49) = 22.57$, $p < .001$; SI Fig. 5.13A). Classifiers also generalized to the main task, showing above-chance evidence for presented stimuli during stimulus onset (7.66%, $t(50) = 13.35$, $p < .001$; SI Fig. 5.13B). Importantly, prior work has demonstrated that sub-second replay can be detected even under substantially reduced classifier accuracy and signal-to-noise conditions, indicating that the comparatively weaker decoding in the MTL does not per se preclude reliable identification of sequential reactivation (e.g., Fig. 5 in Wittkuhn and Schuck 2021, SI in Schuck and Niv 2019).

The MTL showed small but consistent sequential replay effects relative to baseline for all sequences, as reflected in a non-zero intercept ($\beta = -0.002$, $z = -2.99$, $p = 0.003$), and did not appear to differentiate between sequence conditions (main effect condition: $\beta = -0.001$ for all sequence conditions; see Fig. 5.5H). Replay in MTL appeared to covary with the strength of visual replay, as evidenced by a mixed-effects model predicting visual replay from MTL replay ($\beta = 0.036$, $z = 2.998$, $p = 0.003$), indicating cross-region coupling in replay. This coupling was consistent across all conditions (all interaction terms with condition $p > 0.2$). MTL replay did not show any of the modulation of reversal phases, non-local value updates, or relation to structure learning that we observed in visual cortex. Hence, while MTL showed above chance reactivation and correlated with visual replay, it seemed less structured.

To ensure the above results do not reflect inflated statistical baselines, we performed the same analysis on a brain region where we did not expect replay to occur, the somatosensory

cortex. No evidence of replay was observed during prolonged ISIs in this region (intercept: $\beta = 0.000$, SE $= 0.001$, $z = 0.096$, $p = 0.923$, 95% CI $[-0.001, 0.001]$; all condition effects $p > 0.43$, coefficients near zero; see Fig. 5.5H), ruling out the potential issue of an inflated chance baseline that could have affected our results. Both the MTL and the visual ROI significantly differed from this baseline (MTL: $\beta = -0.003$, $z = -3.448$, $p = 0.001$; visual: $\beta = -0.006$, $z = -7.879$, $p < 0.001$).

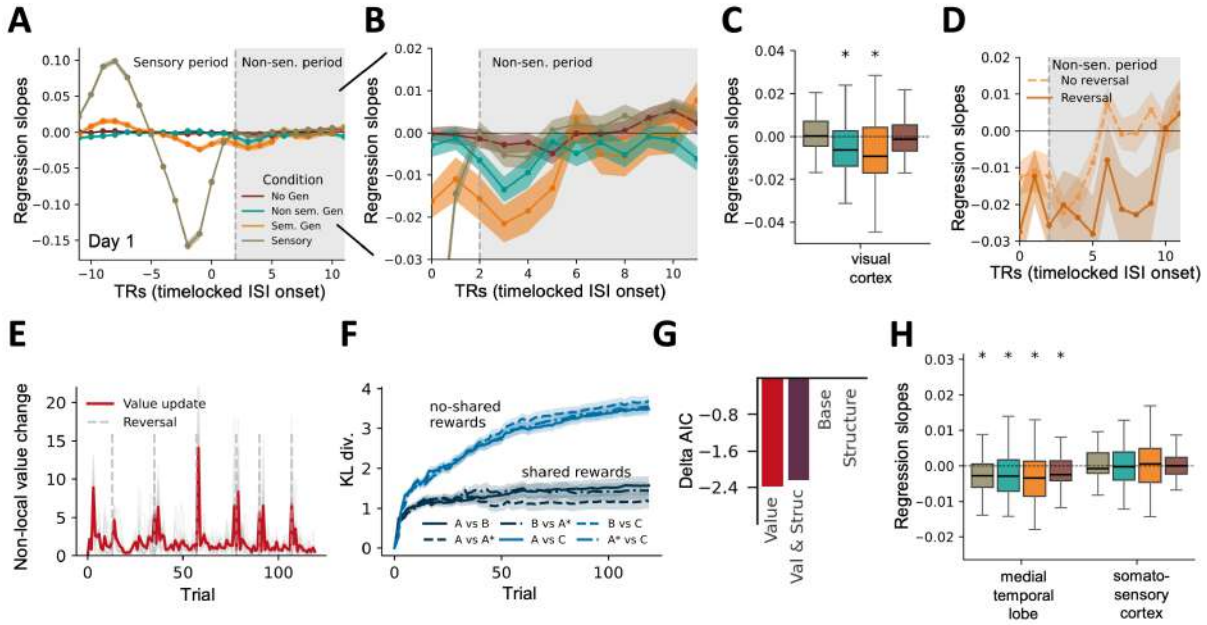

Figure 5.5: **Detecting and relating replay to value updating and structure learning** (A) Regression slope timecourses. Sensory sequences show forward then backward patterns during stimulus presentation. After the sensory effect subsides (gray shaded area), reward-linked sequences (semantic, non-semantic) exhibit selective backward replay, while unrelated sequences do not. (B) Zoomed regression slope timecourses during the non-sensory period, highlighting distinct reactivation for semantic and non-semantic generalization. (C) Aggregated regression slopes during the non-sensory period. Semantic (orange) and non-semantic (turquoise) sequences show selective backward replay. Error bars show SEM. (D) Semantic generalization slopes split by reversal vs. non-reversal trials. Stronger reactivation occurs following reversals, consistent with need for non-local value updating. (E) Latent cause model–derived nonlocal value updating for unobserved options peaks around reward reversal trials. The red line shows the mean trialwise change in unobserved sequence values, averaged across participants with the same reward schedule. The predictor used in the linear mixed-effects model was computed at the subject, trial, and sequence level, capturing the model-derived value update for each sequence (for example, the update to B after observing A on that trial) and relating the magnitude of this update to the amount of reactivation evidence observed for the corresponding sequence. (F) Latent cause model structure learning, quantified as the evolving similarity between paths over time (symmetric KL divergence between their latent-cause probability distributions). Paths that shared rewards showed increasing similarity (decreasing KL divergence), whereas non-shared paths remained distinct. The corresponding predictor in the linear mixed-effects model was the trialwise change in this KL divergence metric, capturing moment-to-moment updates in inferred task structure. (G) Trial-by-trial updating of value predicts replay reactivation beyond a baseline model including only condition as a predictor. (H) Aggregated regression slopes during the non-sensory period for MTL and somatosensory cortex. In the MTL, all conditions show significant backward replay across sequences, albeit weaker than in visual cortex. In contrast, the somatosensory cortex (control) shows no evidence of systematic replay for any condition.

5.2.6 Replay adapts to changes in reward structure

Our results on Day 1 suggested that replay was more tightly coupled with non-local value learning than structure learning. To assess the sensitivity of replay to structure and change more directly, participants returned for a second session on the following day, during which we manipulated their knowledge of the latent structure and the reward correlations, while keeping all other task aspects identical. To ensure that all participants began Day 2 with the same understanding of the structure learned on Day 1, we explicitly revealed which sequences shared the same reward pairing at the end of Day 1. Before the structure was revealed, participants completed a questionnaire indicating which sequences they believed shared a reward level (i.e., yielded a similar number of coins). Nearly half of the participants (49%) identified the correct set of sequences immediately. After receiving a hint that three paths shared the same reward level, an additional 37% selected the correct paths. The remaining 14% responded incorrectly twice.

On Day 2, participants then began by completing two blocks of the task with the previously revealed reward structure from Day 1. As expected, this led to near-ceiling performance in reversal-locked trials (97%, ± 0.012; Fig. 5.6B, C, D), and reduced the previously observed difference between semantic and non-semantic generalization ($t(49) = 1.35$, $p = 0.18$), suggesting participants fully relied on the task structure. After block 2, participants were informed that the underlying graph structure would change in the upcoming block. No information about how it would change was provided, such that participants had to discover the new reward structure through experience. The main difference of the new structure was that the A* path was now coupled with latent reward R2 instead of R1 (Fig. 5.6A). This meant that the A* path led to the same outcomes as C, instead of sharing outcomes with A and B as before, and changed the original generalization patterns in two ways: First, some of the generalizations participants had learned on Day 1 were not valid anymore (e.g. A* ⇒B, A⇒A*, henceforth *No Gen (New)*, dashed brown arrows in Fig. 5.6A); second, other generalizations that were invalid on Day 1 now became valid (A*⇒C, C⇒A*, henceforth *Gen (New)*, green in Figs 5.6A, F–H). Our main question was how this change in task structure led to adaptations in behavior and replay. Specifically, we sought to compare generalization and non-generalization conditions that were consistent with the task structure on Day 1 (*Gen (old)* and *No Gen (old)*, respectively) with those consistent with the new task structure (*Gen(new)* and *NoGen(new)*). After completing the fourth block, the new structure was revealed to participants, allowing them to complete the final block with explicit knowledge of the updated reward associations.

The structure change elicited a drop in accuracy, with some variability across participants (Fig. 5.6C; paired t-test, block 2 vs. block 3: $t(49) = 7.99$, $p < 0.001$). Despite this disruption, participants adapted quickly: performance remained well above chance throughout and improved from 80.6% immediately after the change to 90.8% by the end of Day 2. The recovery following the structure change indicates that the transient drop did not reflect increased diffi-

culty of the new structure but instead participants' relearning of the reward associations (effect of block after the structure change in the LME: $\beta = 0.051$, $z = 5.421$, $p < 0.001$). Reversal-trial performance closely mirrored this overall pattern: accuracy dropped from 95% to 83% after the structure change (paired t-test, block 2 vs. block 3: $t(49) = 3.22$, $p = 0.002$). Participants remained well above chance and showed a slight improvement across blocks, reaching 88% reversal accuracy by block 5, although this increase was not statistically significant (paired t-test, block 3 vs. block 5: $t(49) = 1.30$, $p = 0.20$). This pattern was also evident in the reversal-locked recovery, showing the shallowest drop in the early phase before the structure change, a larger and slower recovery immediately afterward, and a return to early-phase recovery levels once the new structure was revealed, albeit requiring one additional trial after the reversal (Fig. 5.6E). The LCI model recapitulated the main features of structure relearning on Day 2 (Fig. 5.6D). When participants encountered an unexpected sequence–reward pairing, the resulting prediction error led the model to infer that a new latent-cause structure was required to account for the observations. This inference initiated the construction of a revised latent structure and drove a gradual reorganization of value estimates around the newly inferred reward-sharing configuration (see SI Fig. 5.9B). Mirroring human behavior, the model showed a transient drop in predictive accuracy following the contingency change and provided a substantially better fit to Day 2 choices than the temporal-difference models (AICc: 39 vs. 70, 59, and 57 for the TD, 1α, and 3α models, respectively).

We next analyzed replay in the visual ROI on Day 2. In the first step, we confirmed that results during the early phase of Day 2, where the task structure was unchanged from Day 1, mirrored those observed on Day 1. Indeed, we again found strong sensory-driven sequential activation, and broadly comparable evidence for selective backward reactivation of reward-sharing sequences in visual cortex: the non-semantic generalization model showed significant negative coefficients ($t(49) = -3.01$, $p = 0.016$), and the semantic generalization effect remained negative but did not reach significance ($t(49) = -1.84$, $p = 0.071$; note that only 2 blocks were included resulting in decreased power); the no-generalization sequence showed no reliable effect ($t(49) = 1.86$, $p = 0.071$). Unexpectedly, we observed a significant positive effect of the sensory sequence, consistent with forward sequentiality ($t(49) = 2.54$, $p = 0.028$; see SI Fig. 5.14).

MTL results in this early phase were also similar to Day 1, showing backward reactivation for all sequence conditions with comparable signal strength: sensory ($t(49) = -3.26$, $p = 0.008$), non-semantic generalization ($t(49) = -2.08$, $p = 0.043$), semantic generalization ($t(49) = -2.86$, $p = 0.008$), and no-generalization ($t(49) = -2.90$, $p = 0.008$). As before, somatosensory cortex showed no significant effects across conditions (all $|t| \leq 2.05$, all $p \geq 0.182$).

Analyzing replay following the structure change, we observed that the required relearning elicited a targeted adaptation of replay. While replay of still-valid Gen(old) associations ($A \rightarrow B$) was predominantly backward as before ($t(49) = -4.90$, $p < 0.001$, dark green in

Fig. 5.6F, G, H), we found evidence for above chance *forward* replay that reflected the newly valid generalizations (Gen(new), light green $A^* \rightarrow C$, $C \rightarrow A^*$, $t(49) = 4.91$, $p < 0.001$), a finding that held when directly comparing the coefficients of the same condition before vs. after structure change ($t(49) = 2.28$, $p = 0.027$; note that the regression coefficient was part of the no-generalization condition (A^* - C) in the early phase, which as we reported above, had a small positive effect already before the structure change; our direct pre vs post comparison accounts for this effect, however). This increase in replay contrasted with the continued absence of replay in the NoGen(old) condition $t(49) = -1.55$, $p = 0.16$, as well as with no evidence of replay for the now-obsolete A^*–A/B association (NoGen(new); $t(49) = -1.76$, $p = 0.14$, dark brown). In sum, these changes suggest a dynamic adaptation of replay structure that co-occurred with behavioral adaptation to changes in generalization structure.

In the MTL, the structure change did not induce a clear directional flip in replay. All conditions showed evidence of backward sequentiality except the Gen(new), which did not show reliable reactivation in either direction. Specifically, we observed significant negative slopes for the sensory condition ($t(49) = -3.11$, $p = 0.008$), the Gen(old) condition ($t(49) = -2.46$, $p = 0.029$), the No Gen(old) condition ($t(49) = -2.32$, $p = 0.031$), and the No Gen(new) condition ($t(49) = -3.47$, $p = 0.006$), whereas the Gen(new) condition did not reach significance ($t(49) = -1.11$, $p = 0.272$). As before, the somatosensory control region did not reveal significant replay evidence in any condition (all $|t| \leq 1.89$, all $p \geq 0.323$).

To ask which factors could explain the replay we again extracted structure learning and value updating metrics from the LCI model for Day 2. Consistent with the previous session, we found no evidence that structure updating predicted replay beyond a baseline model including condition only (AIC = -7871.5 vs. -7874.6; likelihood ratio test: $\chi^2(2) = 4.956$, $p = 0.292$). Unlike on Day 1, reversal periods per se were not a significant predictor of replay (AIC = -7873.0; likelihood ratio test: $\chi^2(2) = 6.479$, $p = 0.166$). Yet, value updating from the LCI model continued to predict replay significantly (AIC = -7877.8; likelihood ratio test: $\chi^2(2) = 11.259$, $p = 0.024$; see Fig. 5.6I). We did not find a reliable relationship between replay strength and behavioral performance on Day 2, likely due to the limited number of trials before the structure change and the shift in replay direction.

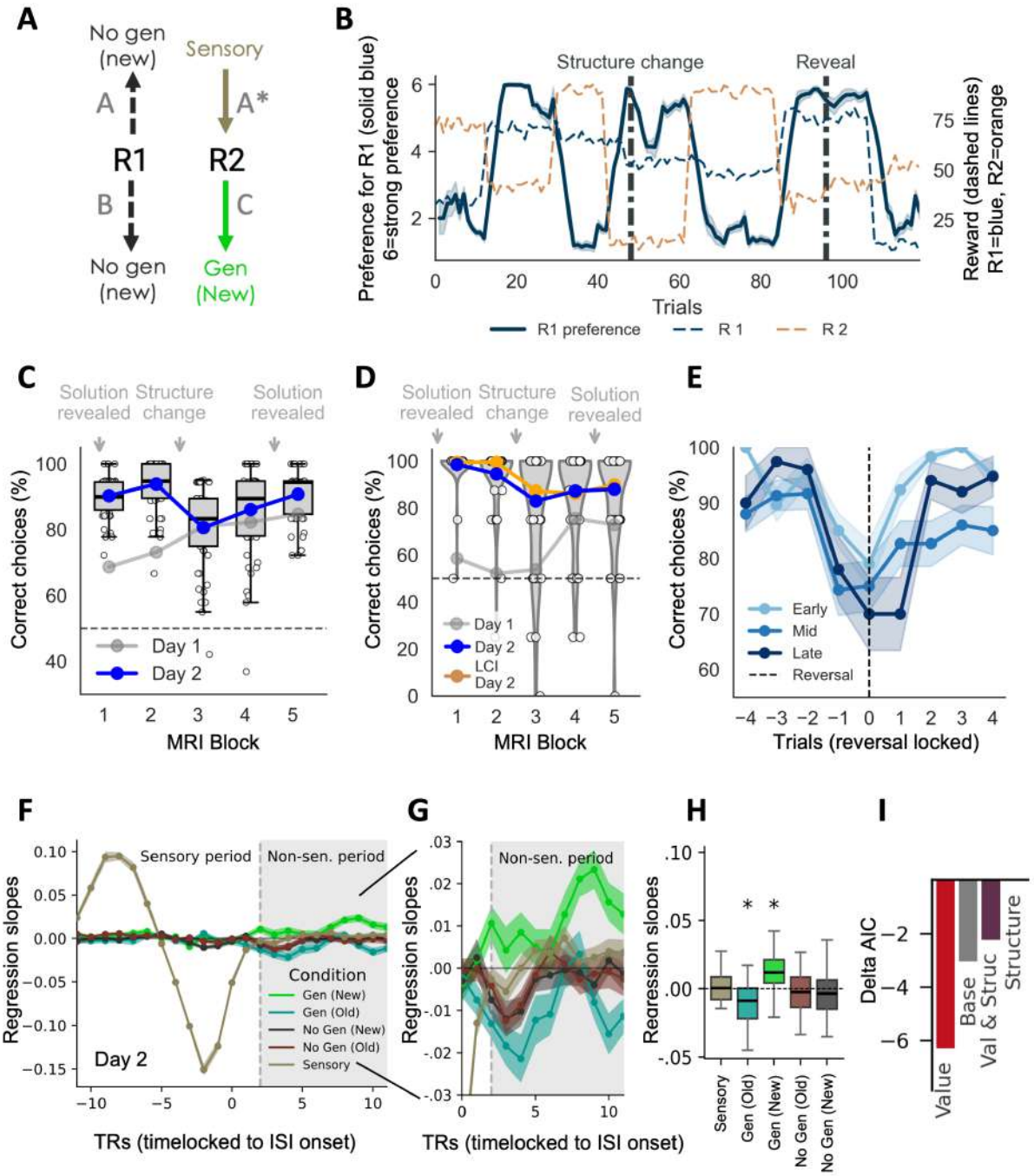
A
No gen (new)
Sensory
A
A*
R1
R2
B
C
No gen (new)
Gen (New)
B
Structure change
Reveal
Preference for R1 (solid blue) 6=strong preference
Reward (dashed lines) R1=blue, R2=orange
Trials
R1 preference
R 1
R 2
C
Solution revealed
Structure change
Solution revealed
Correct choices (%)
Day 1
Day 2
MRI Block
D
Solution revealed
Structure change
Solution revealed
Correct choices (%)
Day 1
Day 2
LCI Day 2
MRI Block
E
Correct choices (%)
Early
Mid
Late
Reversal
Trials (reversal locked)
F
Sensory period
Non-sen. period
Regression slopes
Condition
Gen (New)
Gen (Old)
No Gen (New)
No Gen (Old)
Sensory
Day 2
TRs (timelocked to ISI onset)
G
Non-sen. period
Regression slopes
TRs (timelocked to ISI onset)
H
Regression slopes
Sensory
Gen (Old)
Gen (New)
No Gen (Old)
No Gen (New)
I
Delta AIC
Value
Base
Val & Struc
Structure

Figure 5.6: **Behavior, LCI, and replay on Day 2.** (A) Updated replay and generalization structure following the reassignment of path A* from R1 to R2. Some pre-existing relationships remain unchanged (e.g., generalization across A and B, Gen old; no generalization between A and C, No Gen old), while others reorganize: the former A–A* link becomes invalid (No Gen new) and a new A*–C link emerges (Gen new). (B) Example choice ratings on Day 2, showing preference for R1 (solid blue) alongside the true reward values (dashed lines). The initial phase follows the old structure, followed by a hidden structure change and a later explicit reveal (vertical dashed lines). (C) Accuracy across MRI blocks on Day 2. Participants perform near ceiling early in the session, show a drop after the hidden structure change, and then successfully relearn the updated structure. Day 1 performance (gray) is shown for comparison. (D) Reversal trial accuracy on Day 2. Performance is close to ceiling before the structure change, with a modest drop immediately afterward, indicating rapid adaptation. Day 1 reversal performance (gray) is shown for reference. The LCI model (orange) captures this pattern, reorganizing its latent structure quickly following the change. (E) Reversal recovery on Day 2, time-locked to reversal onset (trial 0). Early reversals occur before the structure change, mid reversals occur after the change but before the reveal, and late reversals occur after the reveal of the new structure. Recovery is fastest during the initial phase; performance after the structure change is slightly reduced but remains above chance, and following the reveal, accuracy returns to near-ceiling within two trials. Note the slight performance dip at trial –1, suggesting that some participants adjusted to emerging reward-level changes before the reversal had fully occurred. (F) Regression slope timecourses on Day 2, showing strong sensory-driven activation and updated replay patterns during the non-sensory period. (G) Zoomed view of the non-sensory period. Sensory and both no-generalization conditions show no reactivation. The old A–B generalization (Gen old) continues to produce robust backward replay, whereas the newly formed A*–C generalization (Gen new) exhibits forward replay. (H) Aggregated regression slopes in the non-sensory period. Both the old (A–B) and new (A*–C) generalization links show significant reactivation; the old link retains backward replay, while the new link exhibits forward replay directionality. (I) Day 2 AIC analysis confirms that trial-by-trial value updating continues to best predict replay reactivation, surpassing a model based on condition alone.

5.2.7 Medial Temporal Lobe Representations Track Structure Learning and Adaptation

Given that generalisation replay relied on structure knowledge, we asked where this task structure was neurally encoded. We hence tested whether visual and MTL representations encoded the reward structure of the task and whether building these representations was associated with replay. We used cross-validated representational dissimilarity matrices (RDMs) capturing the Mahalanobis distances of fMRI activity patterns between all 16 stimuli for each phase (early, late) of each day. The empirical RDMs were compared then to a model RDM that reflected the expected similarities for brain regions that encoded shared reward contingencies, such that stimuli that led to the same reward (A, A*, B on Day 1) were assigned low dissimilarity, and stimuli with different rewards (C vs. others) high dissimilarity (see Fig. 5.7A). To ensure specificity, we set up a number of control RDMs that captured higher similarity among items from the same path (e.g. A1 and A2 are more similar than A1 and B2), similarity of items from the same category, and items that appeared at the same sequence position (see Methods for details). We then orthogonalized the shared reward model RDM with respect to all of these alternative model RDMs, such that any effect could be uniquely attributed to learned task structure.

Given previous evidence [Behrens et al., 2018], we first hypothesized that MTL representations would increasingly align with the shared reward model during structure learning on Day 1, remain high when participants began Day 2 with explicit structure knowledge, and decrease after the Day 2 structure change. A mixed-effects model with Day (Day 1, Day 2), phase (early,

late), model RDM, and their interactions as predictors revealed the predicted three-way interaction (day $\times$ phase $\times$ model RDM: $z = -2.81$, $p = 0.005$), indicating that the relationship between MTL representations and the reward structure changed dynamically across learning and adaptation. Follow-up comparisons confirmed the hypothesized pattern. On Day 1, MTL alignment with the shared reward structure increased from early ($\beta = -0.001$, $z = -0.05$, $p = 0.96$) to late blocks ($\beta = 0.054$, $z = 2.86$, $p = 0.004$), paralleling behavioral acquisition of the structure (Fig. 5.7B). On Day 2, alignment was already high in early blocks ($\beta = 0.047$, $z = 2.55$, $p = 0.011$), reflecting explicit knowledge, but dropped to non-significant levels after the structure change ($\beta = -0.006$, $z = -0.29$, $p = 0.77$), consistent with representational reorganization during adaptation. Further splitting the empirical RDMs into individual sequence pairs and testing late-early differences revealed small but consistent effects: The only nominally significant effect of similarity decreases on Day 2 was between A and A*, i.e. precisely the pair of arms that was changed ($p = 0.112$, $t(45) = 2.14$, $p = 0.038$). This effect, however, did not survive multiple-comparisons correction (all pairs $p_{\mathrm{FDR}} \geq 0.40$).

The same mixed effect model in the visual ROI showed a significant effect for the structure model RDM ($\beta = 1.346$, $z = 4.73$, < 0.001). However, we found no significant interaction of the model with task phase (early vs. late: ($\beta = -0.206$, $z = -0.512$, < 0.609) or day ($\beta = -0.116$, $z = -0.288$, < 0.773) corresponding to the acquisition during day 1 or the adaptation during day 2. We did not test alignment with the new structure on Day 2 because cross-validated RDMs need to be estimated across blocks. The new structure was valid beginning with block 3, and revealed already after block 4. Hence, cross validation would be based on very little data and would require combining one block with explicit knowledge and one with inferred knowledge, resulting in insufficient power and limited interpretability. We found no significant relationship between changes in representational similarity in MTL and visual or MTL replay activity. Models predicting later similarity change from early replay, and models predicting later replay from representational change, were both non-significant (all $p > 0.18$).

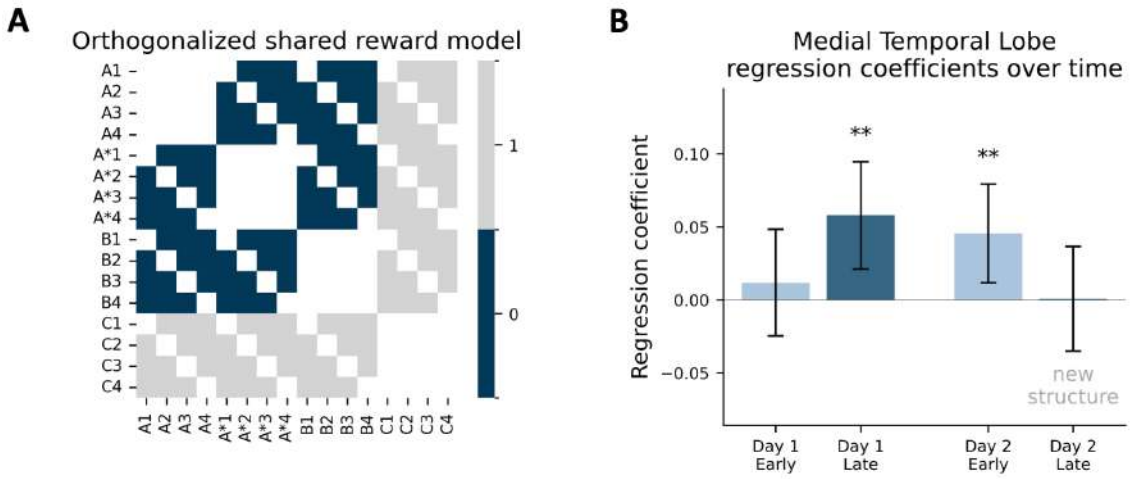

Figure 5.7: **MTL representations track and reorganize according to shared reward structure.** (A) Model representational dissimilarity matrix (RDM) encoding shared reward structure. Stimuli from sequences A, A*, and B (leading to R1) are modeled as similar (dark blue), while comparisons to sequence C (leading to R2) are modeled as dissimilar (light gray). The model was orthogonalized against within-sequence similarity, category similarity, and sequence position to isolate reward structure specifically. (B) Regression coefficients showing alignment between empirical MTL representations and the shared reward model across phases. During Day 1, correlations increase from early to late blocks, indicating gradual acquisition of abstract reward structure. On Day 2, alignment is initially maintained but decreases after the structure change when A* switches from R1 to R2, reflecting representational reorganization. Error bars denote 95% confidence intervals of the regression coefficient estimates.

5.3 Discussion

Adaptive decision-making requires the ability to discover latent structure, generalize knowledge to unobserved situations, and flexibly reorganize behavior when that structure changes. Building on computational accounts in which latent-cause inference underlies the discovery of hidden structure from experience [Gershman et al., 2010, Mirea et al., 2024], we apply this framework to model how humans infer latent structure from trial-and-error feedback and use it to flexibly generalize across shared rewards. We also report that the LCI model outperformed classic reinforcement learning models and captured individual differences in structure learning. Neurally, we find evidence for 'non-local' value propagation as previously reported with MEG [Liu et al., 2021b], albeit we found it localized in visual cortex, rather than in MTL as suggested by Liu et al. [2021b]. Another main novel observation is that replay deeply interacts with latent cause inference, adjusting on a trial by trial basis which experiences are replayed in proportion to the required non-local value updates implemented in a latent cause inference process. Previous work on latent-cause inference has emphasized how experiences separate into distinct latent causes when contingencies change. New latent causes are created when prediction errors signal that the current generative model no longer explains observations, leading to context-specific learning and the preservation of past associations [Gershman et al., 2010]. This state-splitting mechanism accounts for phenomena such as renewal after extinction

or context-dependent memory retrieval. In contrast, our approach highlights a complementary functioning of latent-cause inference: generalization by assigning different observations to a shared latent cause, similar to Mirea et al. [2024]. By mapping distinct experiences onto the same underlying latent cause, the model determines when knowledge learned in one setting can be transferred to another. Thus, rather than proliferating latent causes to explain differences, our account relies on their shared use to integrate information across situations that share an underlying relational structure. This emphasis on generalization links to other representation learning frameworks supporting generalization. The successor representation (SR) [Dayan, 1993, Gershman, 2018, Momennejad et al., 2017, Kahn et al., 2025, Keck et al., 2025] encodes predictive relations among states, forming a cached map of expected future state occupancies that enables rapid reward revaluation but offers limited flexibility when transitions change (but see Wittkuhn et al. [2025] and Sagiv et al. [2025] for evidence linking replay to SR learning and re-learning). The linear reinforcement-learning framework and its default representation (DR) [Piray and Daw, 2021] extend this idea to a stable, policy-independent predictive map that can be reused across changing goals or environments without relearning transition dynamics. Both SR and DR capture structure through prediction rather than inference: by encoding how states unfold through learned transitions, they support generalization by leveraging predictive relations. In contrast, latent-cause inference generalizes by clustering observations under shared hidden causes, allowing the agent to infer when superficially different experiences reflect the same latent structure.

Neurally, we observed two key phenomena that provide insight into how latent structure discovery and neural replay support flexible generalization. First, backward replay in visual cortex selectively reinstated unobserved sequences that shared a latent reward, consistent with a mechanism for non-local value updating [e.g. Liu et al., 2021b] and with the occurrence of learning-related replay in visual cortex [Wittkuhn et al., 2025]. Second, representational similarity analyses revealed that multivariate patterns in the MTL gradually aligned with the abstract reward structure, reorganizing when contingencies changed on Day 2.

During the early sensory period, we observed a coactivation of semantically related reward-sharing sequences, exceeding a baseline derived from the localizer and potentially facilitating value transfer between associated items in line with previous results by Wimmer and Shohamy [2012]. Following the sensory period after the sensory effect subsided reactivation in visual cortex was selectively expressed for reward-associated sequences, ruling out visually driven or nonspecific reactivation. Replay strength increased across Day 1 as participants discovered the underlying structure, peaked after reward reversals, and predicted performance over a longer behavioral horizon. Different previous accounts disagree about the extent to which replay is hypothesized to support value vs. structure learning [Mattar and Daw, 2018, Sagiv et al., 2025]. In the current study, trial-wise fluctuations in replay were best explained by non-local value updates derived from the LCI model rather than by a structure-learning metric, indicating a tight coupling between inferred latent structure and value propagation. Although negative

results should be interpreted with caution — and the current study does not ideally distinguish value from structure learning, which are largely driven by the same reversal events — the lack of evidence for replay involvement in structure learning is striking.

The MTL showed a distinct pattern from visual cortex. While multivariate patterns in the MTL dynamically tracked the reward structure, increasing alignment during Day 1 learning and reorganizing after the Day 2 structure change, MTL replay itself was broad and lacked the selectivity observed in visual areas. This contrasts with previous human replay research utilizing MEG [e.g. Wimmer et al., 2023, Liu et al., 2019, Nour et al., 2021], where we observe different signatures in cortical versus medial temporal regions. Previous simulations confirm that replay can be detected despite reduced SNR [Wittkuhn and Schuck, 2021], and MRI studies have detected replay in the hippocampus [Schuck and Niv, 2019], further validating our MTL findings. Moreover, correlation between visual cortex and MTL replay and the selective presence of reactivation in MTL alongside the absence of replay in a somatosensory control region indicate that the broad MTL replay reflects genuine, albeit less targeted, reactivation. Two interpretations, not mutually exclusive, could explain this regional difference. First, classification in MTL was constrained by modest decoding accuracy, potentially limiting the granularity with which sequence-specific replay can be resolved. Second, MTL replay may reflect more global sampling, providing a structural scaffold over which selective cortical replay operates. This interpretation aligns with Kurth-Nelson et al. [2016], who suggested that MEG-detected replay is driven by sensory regions reflecting MTL-orchestrated reactivation. Under this view, the MTL maintains the abstract task structure and drives the selective replay we observe in cortex.

The structural reversal on Day 2 further revealed how replay tracks the evolving task structure. While backward replay along previously valid generalization links persisted, newly valid links exhibited forward replay. Notably, value updating from our LCI model remained the strongest predictor of replay during relearning. Whether this directional shift reflects replay passively following an already updated representation or actively participating in structural relearning remains an open question.

Together, these findings indicate that latent-structure discovery and neural replay jointly support flexible generalization. Latent causes organize experience into shared reward sources, enabling rapid inference, while replay propagates value across this inferred structure to support non-local generalization. Two key questions remain. First, our reversal manipulation closely coupled structural inference with non-local value updating, making it difficult to isolate their respective contributions. Future work should experimentally dissociate changes in latent structure from changes in value, and track how replay direction evolves as new structure is acquired versus merely exploited. Second, the broad and relatively non-selective replay observed in the MTL remains only partly understood. Clarifying whether this pattern reflects methodological limits on decoding or a coordination signal that scaffolds more selective cortical replay will be essential for understanding how replay and latent structure interact to enable rapid, flexible

behavior beyond direct experience.

5.4 Methods

5.4.1 Participants

Fifty-two neurotypical participants (mean age = 27.4 years, SD = 5.6; 25 females) were recruited from the participant database of the Max-Planck Institute for Human Cognitive and Brain Sciences, Leipzig. All participants had no counter indications for MRI scanning, no history of psychiatric conditions, normal or corrected-to-normal vision and provided written informed consent before participation. The study was approved by the ethics committee at the University of Leipzig (002/22-ek). One participant discontinued during the localizer task on day 1 and was excluded entirely. Another participant did not complete day 2 due to illness, and was therefore only included in day 1 analysis. Finally, 3 participants had incomplete scanning blocks during day 1 due to technical malfunctions. In these cases we use all available data, i.e. only excluded blocks were necessary. This resulted in 51 participants for day 1 analyses and 50 participants for day 2 analyses. Participants received €12 per hour (€96 total for 8 hours across two 4-hour sessions) plus a performance-based bonus up to €34, calculated from points collected during the main tasks and accuracy in localizer tasks.

5.4.2 Experimental Design

The experiment consisted of four fMRI sessions split across two consecutive days. Each day began with a localizer session for classifier training (sessions 1 and 3; 70 minutes), followed by the main task (session 2 and 4; 65 minutes). Both days included a break of approximately 20 minutes between sessions.

5.4.2.1 Sessions 1 and 3 - localizer

Session 1 (Day 1 Localizer): Participants viewed 17 stimuli in a pseudorandom order while performing a 1-back task. Stimuli consisted of 1 reward image (treasure chest) and 16 unique 750ms video clips of everyday objects, including animals (giraffe, zebra), vehicles (sports car, generic car), buildings (modern house, old house), and food items (raspberries, grapes). Videos were selected to create clear categorical groupings, relevant for the main task, while maintaining visual distinctiveness for classification. Stimuli were presented using PsychoPy (version 2022.2.4) and projected onto a screen visible via a mirror mounted on the head coil. Each stimulus was presented 34 times across six blocks with balanced transitions (579 trials in total). The mean inter stimulus interval was 3s sampled from a truncated exponential distribution ranging from 2 to 7s. On occasional probe trials (30% of trials), participants had to identify the stimulus shown immediately before the probe by a left or right button press. Three types of probe trials

were included to ensure attentiveness: (1) the same stimulus and a foil, (2) a different exemplar of the same object category (e.g., a different zebra) and a foil, and (3) two foils, requiring a special button press (down) when neither matched the previously seen object. The maximum response time was 2 seconds. This localizer session provided unbiased training data for the multivariate classifiers. The two localizer sessions across Days 1 and 2 were identical in design. However, participants differed in their knowledge of the task. During the first localizer on Day 1, participants were naive to both the upcoming main task and the latent reward structure. On Day 2, participants again completed the same localizer, but now with full knowledge of the latent structure and the associations among stimuli, although the structure knowledge was not required to perform the localizer task.

5.4.2.2 Sessions 2 and 4 - main task

The main task involved four sequential "bandits" (A, A*, B, and C), each consisting of a fixed sequence of four 750 ms video clips that culminated in a reward outcome. On Day 1, sequences A, A*, and B were linked to a shared latent reward (R1), whereas sequence C was associated with an independent reward (R2). Rewards were matched in accumulated value across all trials. Observed rewards ranged between 0 and 100 and were stochastically sampled from the true mean of R1 or R2 on each trial with a small standard deviation (SD = 2). Throughout the session, true reward means were organized into stable phases punctuated by six reversals (15–20 trials apart). During reversals, mean rewards changed abruptly, typically passing through 1–2 transitional values before stabilizing at the new level. Each participant was assigned to one of four reward schedules (different schedules each day) generated by counterbalancing (i) the order of reward changes and (ii) which of the two latent reward timecourses was assigned to R1 vs. R2. This ensured that group-level effects could be attributed to learning rather than idiosyncratic reward trajectories. Although reversals of R1 and R2 sometimes co-occurred (see Fig. 5.2A/B), reward time-courses were designed to also include phases in which one reward changed while the other remained stable, as well as intervals of gradual drift and abrupt shifts (average correlation between R1 and R2 time courses: r = -0.18). Assignment of video identities to sequence positions was randomized across schedules.

To provide an additional scaffold for structure learning, sequences A and A* also shared semantic category membership at corresponding positions (e.g., both showing animals at position 1, houses at position 2, etc.; the category ordering varied across participants), thereby creating a "semantic generalization" condition (shared reward and shared category), in addition to "non-semantic generalization" (shared reward only). For example, A: giraffe → modern house → raspberries → sports car; A*: zebra → old house → grapes → generic car. The first day concluded with explicit instructions about the latent reward structure.

Day 1 main task (Session 2) comprised 120 trials arranged in 5 blocks, with each sequence shown six times per block. Each trial consisted of a sequential presentation of four videos

(750 ms each) followed by a reward image (750 ms) depicting coins and overlaid with the associated numerical outcome. After observing each sequence and its outcome, participants made preference judgments between two sequence elements (4 s response window). Participants were instructed to choose the option they believed was associated with the more valuable reward at that moment. Confidence was indicated on a 6-point scale, with extreme ratings (1 or 6) reflecting high confidence and values closer to the center indicating greater uncertainty. 100 such judgments were collected over the session, always matching positions within the sequence; positions 2 and 3 were probed more frequently, with probe frequencies of 6, 45, 30, and 9 trials for positions 1–4, respectively. Ten trials involved options that originated from the same reward source and therefore lacked a correct choice under the initial structure (e.g., choices between A and A*, which only acquired different values after the structure change). On these trials, participants used the slider less decisively, showing more center-position ratings rather than strong preferences at the extremes, as confirmed by a linear mixed-effects model predicting ratings based on whether the probed options came from shared versus separate reward sources. Five trials following each reward reversal served as critical tests of latent structure knowledge, requiring participants to generalize value from an observed sequence to an unobserved one. For example, after a reversal, the first four trials might only present sequences A and C, while the preference judgments probe A* and B. On the fifth trial, either A* or B is shown and the remaining unobserved option is probed. Successfully performing these trials therefore requires inferring the relative value of unobserved sequences based on the most recent observed outcome. Each day participants solved 6 reversals which were distributed across the 5 blocks, with blocks 2 or 4 containing 2 reversals.

The reward schedules consisted of stable phases with small fluctuations, gradual drifts, and occasional abrupt changes, creating instances in which reversal choices could be solved based on reward memory. These trials still required adapting reward preferences (i.e., switching from preferring R1 to R2 or vice versa) but could be guided by previously observed outcomes if the drifting reward source had not yet changed substantially. For example, when the shared reward R1 drifted gradually while R2 abruptly dropped from high to low, observing the new R2 outcome together with one R1 branch provided sufficient information to infer the values of the remaining (unprobed) R1 branches. In such cases, only minor value updating through generalization was needed, allowing participants to rely on prior experience. When focusing on trials that required updating reward preferences but could not be guided by previously observed outcomes (e.g., when both rewards changed simultaneously), learning improvements from early (Blocks 1–2) to late (Blocks 4–5) stages were comparable to those observed across all reversal trials. The session concluded with explicit feedback about the latent reward structure.

On Day 2 main task (Session 4), participants first received a reminder of the latent reward structure and completed a short behavioral training block outside the scanner to practice reward-based generalization. Inside the scanner, Blocks 1–2 then assessed performance with full structure knowledge. After Block 2, participants were informed that the underlying reward

structure would change, but were not told how. In Blocks 3–4, the reward contingency of A* switched from R1 to R2, requiring participants to update their generalization strategy. Before Block 5, the new latent structure was explicitly revealed. This manipulation required participants to abandon previously learned semantic associations and base generalization exclusively on the new reward contingencies. Inter-stimulus intervals (ISIs) were drawn from a truncated exponential distribution (mean = 3 s, range 2–7 s). On a subset of trials, the ISI following reward was prolonged by an additional 15 s (average 18 s total) before the preference judgment. These prolonged ISIs occurred eight times per block (40 per session) and were concentrated around reward reversals to provide periods without sensory input during which neural replay was expected to support non-local value updating.

5.4.3 Behavioral Analysis

Performance was quantified using multiple metrics. Overall accuracy was defined as the percentage of trials in which participants selected the option leading to the higher reward. To assess generalization after reversals, we analyzed the accuracy of post-reversal choices involving bandits whose outcomes had not yet been observed since the reversal (i.e., choices requiring generalization from observed to unobserved sequences). Learning dynamics were further characterized by computing trial-wise accuracy time-locked to reward reversals (Fig. 5.6E).

To examine the influence of category information, we analyzed how the category of the observed sequence affected subsequent choices on reversal trials. Specifically, we compared generalization performance for within-category generalization (e.g., A→A* or A*→A) versus across-category generalization (e.g., B→A*/A or A*/A→B) to test whether semantic similarity between observed and generalized sequences facilitated structure discovery (Fig. 5.2E).

Statistical analyses were conducted using linear mixed-effects models implemented in Python 3.11.7 using the package statsmodels (version 0.14.4). Models included a random intercept for each participant and tested the fixed effects of interest. Statistical tests were done using scipy (version 1.12.0). Reversal performance analyses (Fig. 5.2C, 5.6D) were corrected for multiple comparisons using false discovery rate correction.

5.4.4 Latent Cause Inference Model

To formalize how participants discover and update latent task structure, we fitted a latent cause inference (LCI) model [e.g. Gershman et al., 2010, Pisupati et al., 2024, Mirea et al., 2024, Berwian et al., 2025] that assigned observations to latent causes using a Chinese Restaurant Process (CRP) prior [Aldous, 2006]. The model assumed that the observed sequence–reward pair (s_t, r_t) on a given trial was generated by an unobserved latent cause with time varying properties, z_t. Each cause was characterized by (i) a categorical distribution over the four sequences and (ii) a reward expectation. The model then used approximate Bayesian inference to infer which cause had generated the observation, and updated the causes correspondingly.

If participants made a choice between two stimuli, without observing any reward, the model retrieved the associated reward expectations given the shown stimuli. The model involved four mechanisms: (1) computation of a prior of latent causes, (2) computation of likelihoods of observed stimuli and reward given each latent cause, (3) sequential Bayesian inference of the posterior using particle filtering, (4) value estimation and decision making. These steps are described in detail below. Model parameters were fitted for each subject by maximum likelihood estimation as described below.

5.4.4.1 CRP Prior

The LCI model maintained a prior over latent causes z_i given the previous latent causes in all preceding trials. The main idea of the CRP prior we used here is that the probabilities of existing causes were proportional to how often they had occurred in previous trials, while the probability to generate new causes was governed by a concentration parameter α and decreased over time:

$$P(z_t = k \mid z_{1:t-1}) = \frac{n_k}{t-1+\alpha} \quad \text{for existing } k, \qquad P(z_t = \text{new}) = \frac{\alpha}{t-1+\alpha}, \tag{5.1}$$

We denote the latent causes on trial t by z_t. Let k index an existing cause, with n_k the number of previous assignments to cause k. Thus, the CRP prior favors reusing existing causes in proportion to how often they have been inferred before, while allowing occasional creation of new causes controlled by the concentration parameter α.

5.4.4.2 Likelihoods

Categorical sequences

Each latent cause maintains a Dirichlet distribution that describes how probable each of the four possible sequences is given each cause. Note that, for simplicity, we do not model the semantic similarity between these sequences, and instead treat them as four distinct categorical observations. At the beginning of the experiment we assume a flat Dirichlet with pseudo-counts $(1, 1, 1, 1)$. Let $n_{k,s}$ denote the number of times sequence s has been assigned to cause k. A category-increment parameter η controls how strongly each observation updates the categorical distribution:

$$\tilde{n}_{k,s} = \eta\, n_{k,s}, \qquad \tilde{N}_k = \sum_s \tilde{n}_{k,s}. \tag{5.2}$$

The categorical probability for existing causes is then

$$P(s_t = s \mid z_t = k) \;=\; \frac{\tilde{n}_{k,s} + 1}{\tilde{N}_k + 4}. \tag{5.3}$$

For a newly created latent cause, we use the uniform categorical prior $P(s_t = s \mid z_t =$

new) $= 1/4$.

Rewards

Reward likelihood is a uniform–Gaussian mixture:

$$P(r_t \mid z_t = k) = h \cdot \mathcal{U}(0, 100) + (1 - h) \cdot \mathcal{N}(\mu_k, \tau^2), \tag{5.4}$$

with hazard rate h [e.g., Nassar et al., 2010]. The uniform component allows abrupt reward changes to be incorporated without necessitating creation of new latent causes, preserving the shared structure needed for value generalization. In the online update used here, μ_k is set to the most recent reward assigned to cause k and a fixed τ^2 of 5 is used for existing causes; for a new cause we use $\mu = 50$ and a broader variance of 100.

5.4.4.3 Computing the Posterior and Particle Filtering

Because exact Bayesian inference of the posterior is non-tractable, we approximate inference using a particle filtering approach. The above process of computing the prior and likelihoods was repeated for each of 100 particles, on each trial [Sanborn et al., 2010]. This process involved the following steps for every particle.

1. Compute the CRP prior over existing causes plus a potential new cause.
2. For each candidate cause, compute the likelihood combining categorical and reward probabilities $P(s_t \mid z_t = k)\, P(r_t \mid z_t = k)$ (for potential new latent causes using uniform sequence and broad reward likelihood)
3. Branch the particle for each candidate cause and assign the weight of each candidate by prior$\times$likelihood. This creates for each particle multiple candidates (one per existing cause and an additional for a new cause).
4. Update the categorical and reward estimate of each candidate based on the current observation.
5. Normalize branch weights across the full candidate pool.
6. Resample back to 100 particles; weights are reset to uniform after resampling.

5.4.4.4 Calculating Value Estimates and Value-Based Decision-Making

After each observation, the particle filter expands and resamples the particle set. Each particle $i \in \{1, \ldots, M\}$ maintains a set of currently instantiated latent causes K_i, each with an associated reward mean μ_{ik} and categorical counts over sequences. The categorical posterior-predictive $P_i(s \mid k)$ for each cause is defined as in the Categorical sequences section.

For a queried sequence s, each particle forms a mixture over its own latent causes by normalizing these categorical predictive probabilities:

$$w_{ik}(s) \;=\; \frac{P_i(s \mid k)}{\sum_{j \in K_i} P_i(s \mid j)} \qquad (\text{so } \sum_{k \in K_i} w_{ik}(s) = 1). \tag{5.5}$$

The particle-level value of sequence s is then the expectation of the reward means under these mixture weights:

$$V_i(s) \;=\; \sum_{k \in K_i} w_{ik}(s)\, \mu_{ik}. \tag{5.6}$$

Because particles are resampled and assigned equal weight after every trial, the model's predicted value is the simple average across particles:

$$V(s) \;=\; \frac{1}{M} \sum_{i=1}^{M} V_i(s). \tag{5.7}$$

Intuitively, each particle represents a distinct hypothesis about how observations cluster into latent causes; $V(s)$ averages these hypotheses to obtain the predicted expected reward for each sequence which builds the basis for subsequent decision-making.

To produce binary choices between two sequences s_1 and s_2, the model first computes the value of each sequence as described above, by averaging each particle's reward expectation weighted by its latent-cause mixture probabilities. These values $V(s_1)$ and $V(s_2)$ are then transformed into choice probabilities using a softmax function with inverse-temperature parameter θ:

$$P(\text{choose } s_1 \text{ over } s_2) = \frac{\exp\big(\theta\, V(s_1)\big)}{\exp\big(\theta\, V(s_1)\big) + \exp\big(\theta\, V(s_2)\big)}. \tag{5.8}$$

5.4.4.5 Parameter Estimation and Model Comparison

Participant-specific parameters were estimated for each day separately by maximizing the likelihood of observed choices using PyBADS [Acerbi and Ma, 2017]. The LCI model had the following 4 free parameters and bounds:

- $\alpha \in [0, 1]$: CRP concentration (new latent cause control)
- $\theta \in [0, 1]$: softmax inverse-temperature (choice stochasticity)
- $h \in [0, 1]$: hazard rate (mixture of uniform and gaussian reward distribution)
- $\eta \in [1, 10]$: category increment (evidence weight per sequence observation)

We used a structured multi-start (3 starts per parameter; $3^4 = 81$ runs) and selected the best log-likelihood. Model comparison used the corrected Akaike Information Criterion (AICc):

$$\text{AICc} = 2k + 2\,\text{NLL} + \frac{2k(k+1)}{n-k-1}, \tag{5.9}$$

computed per participant.

To elucidate the potential computational functions of replay, we derived two trialwise metrics from the fitted latent cause inference (LCI) model. The first metric captured nonlocal value updating, quantifying the extent to which observing a reward led to changes in the inferred values of sequences that were not directly experienced on that trial. This measure was computed on a sequence- and trial-specific level, allowing us to link replay-related neural activity to the model's propagation of reward information across related task elements (Fig. 5.5E). The second metric reflected structure learning, tracking how the model inferred latent task structure over time. For each trial, we obtained the probability distribution over latent causes for every sequence and quantified their pairwise similarity by computing symmetric Kullback–Leibler (KL) divergences between these distributions. Smaller KL divergences indicated increasing overlap in latent-cause representations, reflecting convergence toward a shared structural understanding independent of reward (Fig. 5.5F). To match the trialwise dynamics of nonlocal value change, we additionally computed the trialwise change in the KL divergence metric, providing a dynamic measure of structural updating across trials. To align model dynamics with the experimental manipulation of task structure on Day 2, the model was initialized with two latent causes, each assigned to the corresponding sequences with a pseudocount of 5. After the first two blocks, the categorical likelihoods and priors were broadened by a fixed decay parameter of 0.8 to simulate increasing uncertainty, and following block 4, the latent-cause assignments were updated to reflect the new post-change reward structure. The resulting trialwise estimates of both nonlocal value updating and structure learning were used as predictors in mixed-effects models linking behavior and replay (Fig. 5.5G).

5.4.4.6 Alternative Temporal Difference Learning Models

To dissociate flexible latent-structure learning from simpler accounts, we compared the LCI model against temporal-difference (TD) models with and without hard-coded generalization. All TD variants track a separate value $V(s)$ for each sequence $s \in \{A, B, A^*, C\}$ and update the value of the experienced sequence s_t on trial t via a standard delta rule,

$$\delta_t = r_t - V(s_t), \qquad V(s_t) \leftarrow V(s_t) + \alpha\,\delta_t, \tag{5.10}$$

where r_t is the observed reward and α a learning rate. In the *no-generalization* baseline, unvisited sequences are never updated, preventing generalization and requiring values to be relearned independently after reversals. In the *hard-coded generalization* variants, the same prediction

error δ_t is also applied to predefined related sequences according to the task's latent structure (e.g., between A and A^*, and between A/A^* and B). One variant uses a single learning rate α for both direct and generalized updates, enforcing fixed-strength generalization from the outset. A second variant uses three separate learning rates α_{exp} for direct experience, α_{semantic} for generalization between A and A^*, and $\alpha_{\text{non-semantic}}$ for generalization between A/A^* and B, allowing graded influence along different generalization types. All TD models translate values into choices using the same softmax choice rule and inverse-temperature θ used by the LCI model. Critically, because their generalization structure is prespecified and static, these models cannot discover latent structure, nor flexibly adapt their structure when reward contingencies change (e.g., on Day 2).

5.4.5 MRI

5.4.5.1 MRI data acquisition and preprocessing

MRI data were acquired using a 32-channel head coil on a 3T Siemens Magnetom Prisma Fit system. Functional MRI (fMRI) scans were collected in a transverse orientation using T2*-weighted gradient-echo echo-planar imaging (GE-EPI) with multiband acceleration, sensitive to BOLD contrast. This sequence achieves high temporal resolution while maintaining good spatial coverage and resolution. We acquired 72 transverse slices with a thickness of 2 mm and an in-plane resolution of 2 × 2 mm. Imaging parameters included a multiband acceleration factor of four, a repetition time (TR) of 1.5 seconds, and an echo time (TE) of 24.2 ms. The field of view (FoV) was 204 mm in both read and phase directions, and slices were collected in an interleaved order. Fat saturation was applied, and the flip angle was set to 70°. The first five volumes of each block were discarded to allow for scanner equilibration. Additionally, each day a high-resolution T1-weighted anatomical scan (1 × 1 × 1 mm) was acquired for spatial normalization and coregistration. To correct for susceptibility-induced distortions, we collected field maps using spin-echo EPI sequences with opposing phase encoding polarities (anterior-posterior and posterior-anterior), which were used for distortion correction. All fMRI preprocessing steps were performed using fMRIPrep.

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 23.1.4 (Esteban et al. [2019]; RRID:SCR_016216), which is based on *Nipype* 1.8.6 (Gorgolewski et al. [2011]; Gorgolewski et al. [2018]; RRID:SCR_002502).

Preprocessing of B0 inhomogeneity mappings A total of 6 fieldmaps were found available within the input BIDS structure for this particular subject. A *B0*-nonuniformity map (or *fieldmap*) was estimated based on two (or more) echo-planar imaging (EPI) references with `topup` (Andersson et al. [2003]; FSL None).

Anatomical data preprocessing A total of 2 T1-weighted (T1w) images were found within the input BIDS dataset. All of them were corrected for intensity non-uniformity (INU)

with `N4BiasFieldCorrection` [Tustison et al., 2010], distributed with ANTs (version unknown) [Avants et al., 2008, RRID:SCR_004757]. The T1w-reference was then skull-stripped with a *Nipype* implementation of the `antsBrainExtraction.sh` workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using `fast` [FSL (version unknown), RRID:SCR_002823, Zhang et al., 2001]. An anatomical T1w-reference map was computed after registration of 2 T1w images (after INU-correction) using `mri_robust_template` [FreeSurfer 7.3.2, Reuter et al., 2010]. Brain surfaces were reconstructed using `recon-all` [FreeSurfer 7.3.2, RRID:SCR_001847, Dale et al., 1999], and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle [RRID:SCR_002438, Klein et al., 2017]. Volume-based spatial normalization to two standard spaces (MNI152NLin6Asym, MNI152NLin2009cAsym) was performed through nonlinear registration with `antsRegistration` (ANTs (version unknown)), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization and accessed with *TemplateFlow* [23.0.0, Ciric et al., 2022]: *FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model* [Evans et al. [2012], RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov et al. [2009], RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym].

Functional data preprocessing For each of the 30 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` [FSL , Jenkinson et al., 2002]. The estimated *fieldmap* was then aligned with rigid-registration to the target EPI (echo-planar imaging) reference run. The field coefficients were mapped on to the reference EPI using the transform. BOLD runs were slice-time corrected to 0.699s (0.5 of slice acquisition range 0s-1.4s) using `3dTshift` from AFNI [Cox and Hyde, 1997, RRID:SCR_005927]. The BOLD reference was then co-registered to the T1w reference using `bbregister` (FreeSurfer) which implements boundary-based registration [Greve and Fischl, 2009]. Co-registration was configured with six degrees of freedom. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. [2014]) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. [2002]). FD

and DVARS are calculated for each functional run, both using their implementations in *Nipype* [following the definitions by Power et al., 2014]. The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction [*CompCor*, Behzadi et al., 2007]. Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, a mask of pixels that likely contain a volume fraction of GM is subtracted from the aCompCor masks. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's *aseg* segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each [Satterthwaite et al., 2013]. Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. Additional nuisance timeseries are calculated by means of principal components analysis of the signal found within a thin band (*crown*) of voxels around the edge of the brain, as proposed by [Patriat et al., 2017]. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNI152NLin6Asym space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): *fsnative*, *fsaverage*. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels [Lanczos, 1964]. Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.10.1 [Abraham et al., 2014, RRID:SCR_001362], mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

5.4.5.2 Copyright Waiver

The above boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts *unchanged*. It is released under the CC0 license.

5.4.5.3 Multivariate Pattern Analysis for Replay Detection

We trained multivariate pattern classifiers to decode individual stimulus representations from fMRI activity patterns, following established methods for detecting fast sequential reactivation [Schuck and Niv, 2019, Wittkuhn and Schuck, 2021, Wittkuhn et al., 2025]. Classifiers were optimized and trained exclusively on task-independent localizer data (Sessions 1 and 3), where all 16 video stimuli were presented with balanced frequencies and transitions. Following Wittkuhn and Schuck [2021], participant-specific anatomical ROIs were defined on the cortical surface using FreeSurfer's automated parcellation of individual T1-weighted reference images obtained during fMRIPrep preprocessing [Dale et al., 1999]. The occipito-temporal ROI included the cuneus (1005/2005), lateral occipital sulcus (1011/2011), pericalcarine gyrus (1021/2021), superior parietal lobule (1029/2029), lingual gyrus (1013/2013), inferior parietal lobule (1008/2008), fusiform gyrus (1007/2007), inferior temporal gyrus (1009/2009), parahippocampal gyrus (1016/2016), and middle temporal gyrus (1015/2015). The medial temporal lobe ROI comprised the hippocampus (17/53), parahippocampal gyrus (1016/2016), and entorhinal cortex (1006/2006). As a control ROI, we included primary somatosensory cortex, defined as the postcentral gyrus (1022/2022).

Within each anatomical ROI, voxel selection and classifier hyperparameters were optimized using a nested cross-validation procedure. Voxel selection was based on two criteria: activation strength and reliability. First, we quantified activation strength as the mean BOLD signal intensity across the entire time course, excluding voxels with poor signal quality or signal dropout.

Second, voxel reliability was assessed using a split-half correlation approach [Tarhan and Konkle, 2020]. Beta estimates were derived from a run-wise first-level GLM implemented in Nilearn (SPM canonical HRF; cosine drift model; high-pass filter = 1/128 Hz; voxel-wise standardization; version 0.10.3). For each run, the design matrix included onsets and durations for all 16 video stimuli (modeled as 0.75 s events) and, when present, a 2 s *probe* condition. Nuisance regressors were added per run and comprised the global signal, framewise displacement, six motion parameters and their temporal derivatives (12 total), and six anatomical CompCor components (aCompCor00–05). Voxel reliability was then computed by correlating beta response profiles between odd (runs 1/3/5) and even (runs 2/4/6) localizer runs. For each voxel,

condition-wise beta maps were averaged within odd and even sets to yield paired response profiles across all regressors of interest. Pearson correlations between these profiles produced voxel-wise reliability maps and allowed the selection of the reliable subset of voxels for subsequent analyses.

Activation and reliability thresholds were determined within the nested cross-validation and applied iteratively to identify the combination yielding optimal classification performance. In visual cortex, best performance was achieved when selecting the 50% most active and 70% most reliable voxels, using training data combined across both localizer sessions (Sessions 1 and 3). In contrast, for the MTL, optimal performance was obtained with more restrictive thresholds, selecting the 30% most active and 30% most reliable voxels. Notably, training the MTL classifier across both sessions decreased performance relative to session-specific training. Thus, for subsequent main-task analyses and replay detection, the MTL classifier was trained on Session 1 only. We implemented one-versus-rest logistic regression classifiers with L2 regularization, chosen for their efficiency with high-dimensional data and probabilistic outputs necessary for replay detection. Inside the nested cross-validation, we additionally optimized spatial smoothing, where best cross-validated classification was obtained with 4mm FWHM spatial smoothing for all regions. We further applied temporal smoothing: For each stimulus, we extracted classifier training data from the TR closest to 5s post-stimulus onset and the preceding TR in the visual cortex. For the MTL, we obtained best classification smoothing over the pre- and succeeding TR (i.e. averaging over TRs 4-6), maximizing signal-to-noise ratio while capturing peak BOLD response. To assess classifier validity, we first examined potential confusions between stimuli. Classifiers assigned only marginally higher probabilities to stimuli within the same category compared to other stimuli, indicating little to no category-level bias (see confusion matrix, Fig. 5.4A). We then validated performance during sequence observation in the main task, where classifiers successfully decoded all 16 stimuli with distinct probability peaks aligned to stimulus presentation times (Fig. 5.4C). One classifier corresponding to a house stimulus was excluded from subsequent replay analyses. During reward presentation (the treasure chest), which was not included in sequence modeling, this classifier showed increased decoding probabilities, suggesting a reward-evoked modulation of its time course that could otherwise be confounded with replay-related reactivation.

5.4.5.4 Sequential Replay Detection

We applied the sequential slope ordering analysis (SODA) from Wittkuhn and Schuck [2021], Wittkuhn et al. [2025] to detect sequential reactivation of stimulus representations during extended rest periods. SODA leverages the temporal dynamics of overlapping hemodynamic responses: when a neural sequence is rapidly replayed, each element evokes a hemodynamic response offset in time. Due to the sluggish nature of the BOLD signal, these overlapping responses create characteristic temporal patterns in classifier probability time courses. The key

insight behind SODA is that probabilistic classifiers applied to overlapping BOLD activity produce systematic fluctuations that preserve information about sequential order. To evaluate the proposed sequentiality for every TR the proposed ordering is regressed onto the probability estimates. If stimuli are replayed in the order $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4$, the resulting classifier probabilities will show ordered dynamics: initially, earlier items (e.g., S_1) will have higher probabilities than later items (e.g., S_4), creating a forward-ordered pattern. As the hemodynamic responses evolve, this pattern reverses, with later items showing higher probabilities, a backward-ordered pattern.

To test whether such ordering exists, we inserted eight extended 18-second intervals per block (40 per session), during which participants viewed only a fixation cross. These 18-second pauses occurred immediately following reward presentation, i.e. before decision prompts and concentrated around reward reversals when non-local updating is critical. We then applied trained classifiers for all 16 stimulus categories to every TR within these intervals, such as to observe spontaneous reactivation without external input during these time windows. To test for sequential reactivation, we categorized potential replay based on relationships to the just-observed sequence into four possible sequential (re)activation events, as described in the main text: Sensory sequential activation, replay that reflects semantic generalization, replay that reflects non-semantic generalization and replay of sequences to which reward information could not be generalized.

For each category and TR, we quantified sequentiality by regressing sequence position (1–4) against classifier probabilities for the four stimuli comprising that sequence. The resulting regression coefficient reflects both the strength and direction of sequential ordering. To distinguish stimulus-driven responses from post-sensory reactivation, we modelled the full sensory response and defined the transition point following the full forward and backward phases of the sequentiality signal as the boundary between "sensory" and "non-sensory" phases. At this point, classifier probabilities for the just presented (sensory) stimuli returned to chance level, and regression coefficients of the sequentiality metric are back at 0. All subsequent effects in the non-sensory phase therefore cannot be attributed to ongoing sensory input and are interpreted as reflecting internally generated sequence reactivation. All analyses focused on full four-item sequences rather than pairwise transitions, differing from alternative approaches focusing on pairwise transitions [Liu et al., 2021a].

5.4.5.5 Representational similarity analysis

We employed representational similarity analysis (RSA) [Kriegeskorte et al., 2008] to examine how neural representations of the 16 task stimuli reorganized to reflect the underlying reward structure during learning and adaptation. RSA quantifies the similarity structure of neural patterns by computing pairwise distances between all conditions, resulting in a representational dissimilarity matrix (RDM). By comparing empirical RDMs to theoretical model RDMs, we

can test specific hypotheses about how information is organized. We constructed a binary model RDM encoding the shared reward structure, where stimuli leading to the same reward show low dissimilarity while those with different rewards show high dissimilarity. Specifically, the model RDM contained:

- Low dissimilarity (0) for stimulus pairs sharing rewards: A-B, A-A*, B-A*, and similarities within sequences
- High dissimilarity (1) for stimulus pairs with different rewards: all comparisons between {A, B, A*} and C

To account for potential confounds, we constructed control RDMs encoding alternative organizational principles: within-sequence similarity (e.g., elements within the A sequence), categorical similarity, and sequential position (all first elements across sequences). The shared reward model was orthogonalized with respect to these control models to ensure effects were specifically driven by reward structure rather than perceptual or sequential factors.

To obtain the empirical RDM, we estimated stimulus-evoked BOLD responses using mass-univariate GLMs implemented in Nilearn. Each of the 5 functional runs per main task session was modeled independently using a standard GLM approach with canonical SPM hemodynamic response function, AR(1) noise model, 4 mm FWHM spatial smoothing, and high-pass filtering (128 s cutoff).

Design matrices included 18 stimulus regressors (one per video stimulus, plus choice and sequence-onset events) convolved with the HRF, and 14 nuisance regressors (global signal, framewise displacement, 6 motion parameters, 6 anatomical CompCor components). Stimuli were modeled as 750 ms events at presentation onset; choice periods used actual reaction times as durations.

For each stimulus, we computed t-statistics contrasting stimulus-specific activation against baseline, yielding 16 t-maps per run. Following the GLM, we computed empirical RDMs based on the multivariate noise-normalized cross-validated Mahalanobis distance [Walther et al., 2016] as implemented in the PyRSA toolbox (https://github.com/Charestlab/pyrsa) [Nili et al., 2014]. We estimated the empirical RDM for each experimental phase (early blocks (1&2) and late blocks (4&5) of each day), resulting in a total of 4 empirical RDMs (early and late for each day). This approach allowed to balance stable parameter estimates using cross-validation, while still resolving the temporal dynamics unfolding over the course of the task, capturing the initial learning phase during day 1 followed by the relearning on the second day.

Based on prior literature, we focused on the MTL as a region often associated with representing abstract task structure. ROIs were defined using the same Freesurfer approach described in the replay section, combined with the activation and reliability thresholds. To test our hypothesis of dynamic structure learning, we employed a linear mixed-effects model testing the three-way interaction:

$$\text{Empirical RDM} \sim \text{Model RDM} \times \text{Day} \times \text{Phase} + (1|\text{Subject}) \tag{5.11}$$

This interaction tests whether representational alignment with the task structure increases from early to late on Day 1 (structure learning) and decreases from early to late on Day 2 (adaptation to structure change)

5.5 Supplementary Information

Remaining behavioral schedules

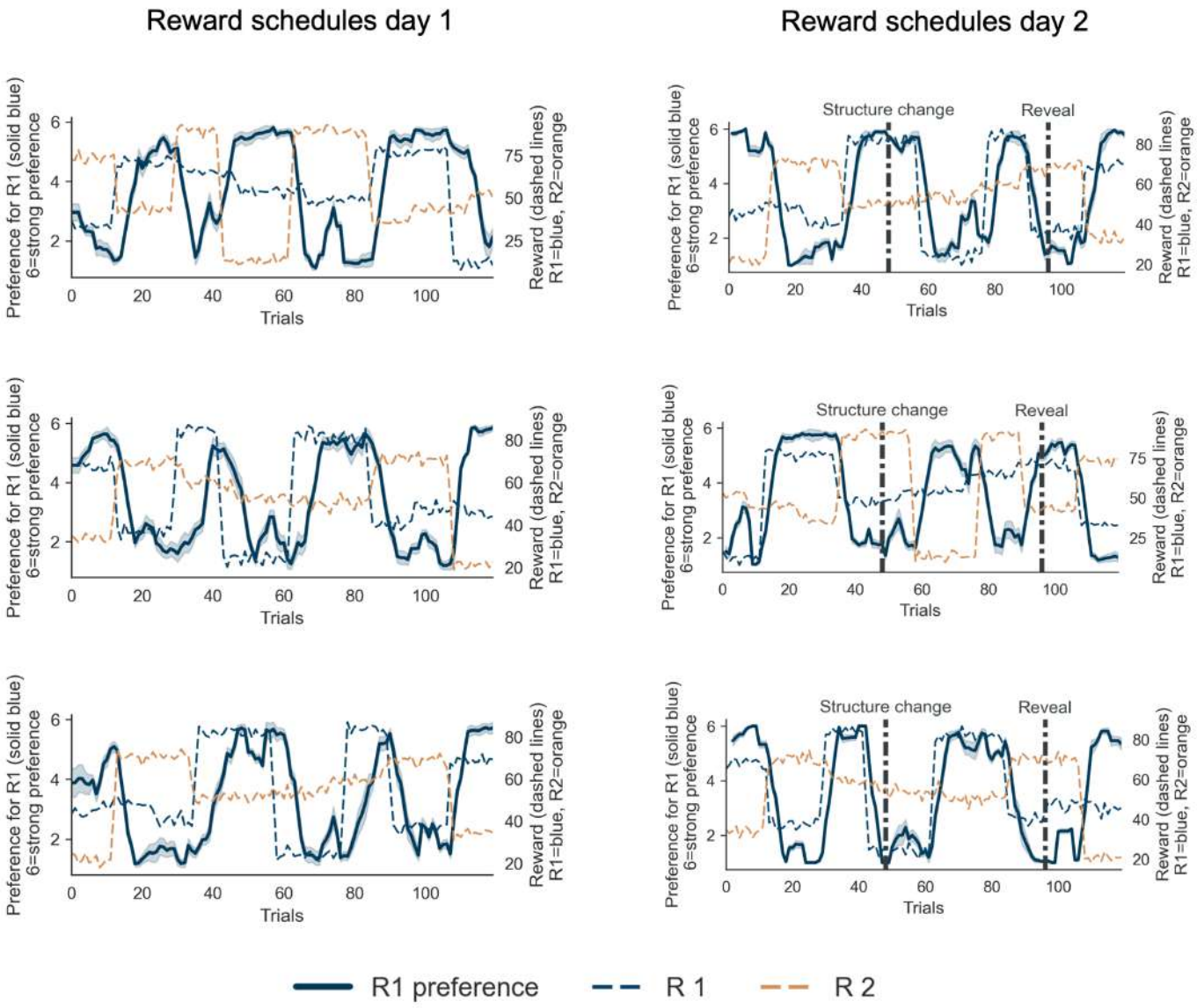

Figure 5.8: **Choice ratings and reward schedules across days.** Left column: the three remaining reward schedules for Day 1, showing trial-wise choice preference (solid blue) alongside the true reward means for R1 and R2 (dashed lines). Right column: the corresponding schedules for Day 2, including the hidden structure change and subsequent explicit reveal.

Latent cause inference model

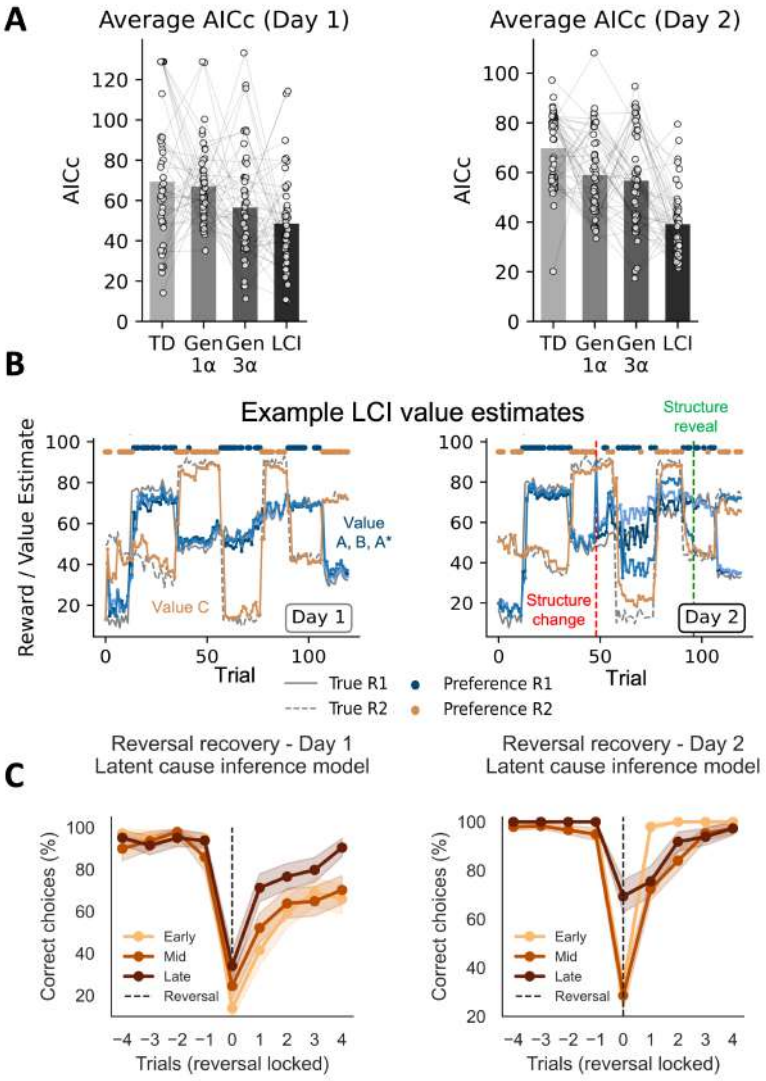

Figure 5.9: **LCI model comparison, latent-cause value trajectories, and reversal adaptation** (A) AICc scores for all models on Day 1 and Day 2. Across both days, the latent cause inference (LCI) model provides the best fit to participants' choices, outperforming both non-generalizing temporal-difference (TD) learning and TD models with hard-coded generalization (Gen 1α, Gen 3α). Dots show individual participants; bars show group means. (B) Example trial-wise value estimates generated by the LCI model. On Day 1, the model converges onto the shared reward structure linking A, A*, and B, while treating C as independent. On Day 2, value estimates reorganize when A* switches reward sources (red line), and stabilize once the new structure is explicitly revealed (green line). Grey traces display the true underlying reward means; colored traces show the model-inferred values. (C) Reversal behavior of the LCI model. The model captures the sharp drop in accuracy at reversals (trial 0) and the subsequent recovery. Recovery accelerates from early to late blocks on Day 1 as structure knowledge improves, and remains robust on Day 2 despite the hidden structure change.

Localizer task session 1 and 3 and stimuli

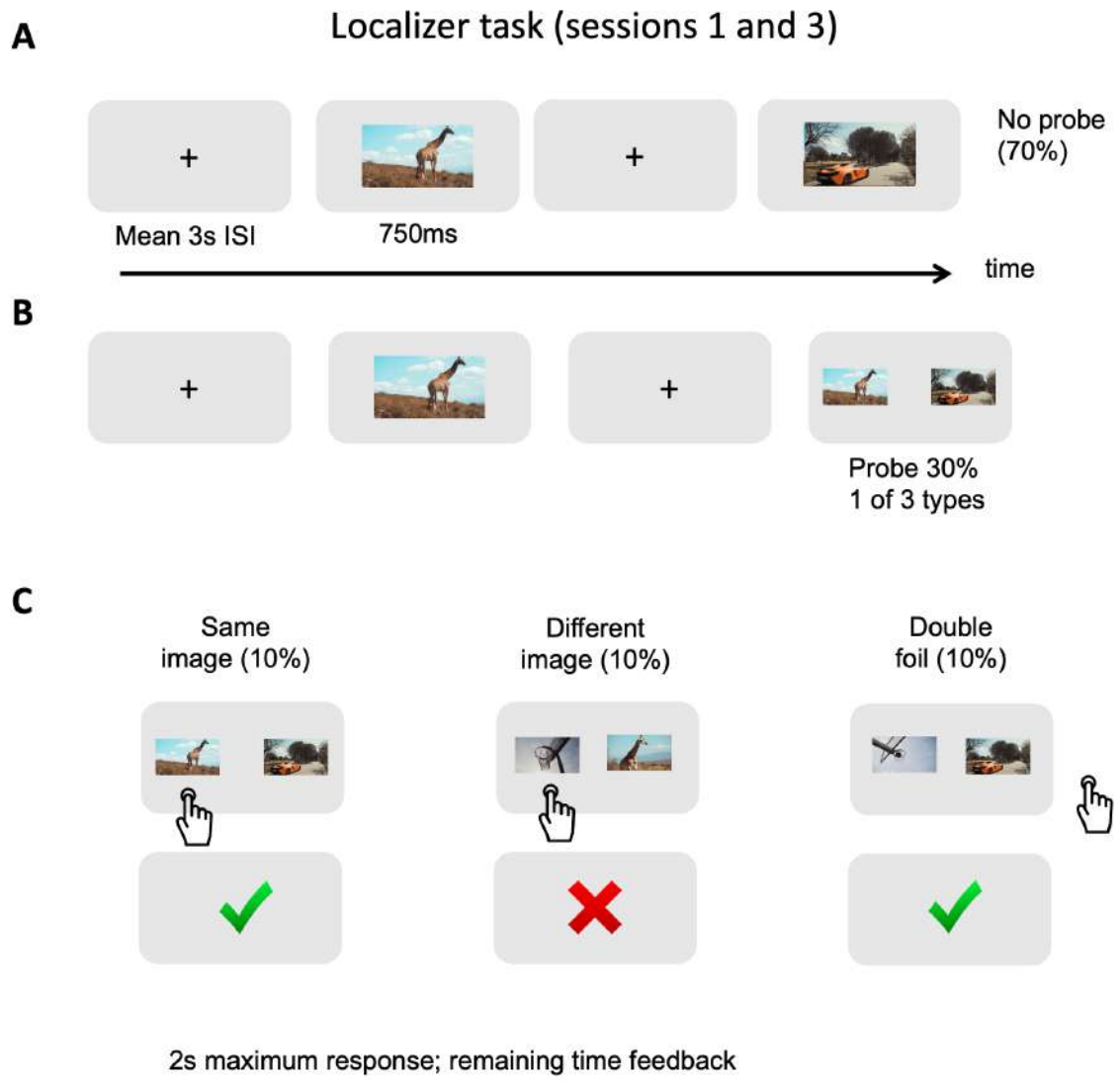

Figure 5.10: **Localizer task (sessions 1 and 3).** (**A**) Trial structure during non-probe trials (70%). Each image was displayed for 750 ms, separated by a mean inter-stimulus interval of 3 s (range: 2–7 s). (**B**) On probe trials (30%), one of three probe types followed the stimulus. (**C**) Probe types required indicating whether: the same exemplar reappeared (same image), a different exemplar from the same category appeared (different image), or neither image matched the preceding stimulus (double foil; correct down-arrow response). Participants had up to 2 s to respond, after which accuracy feedback was shown for the remainder of the interval.

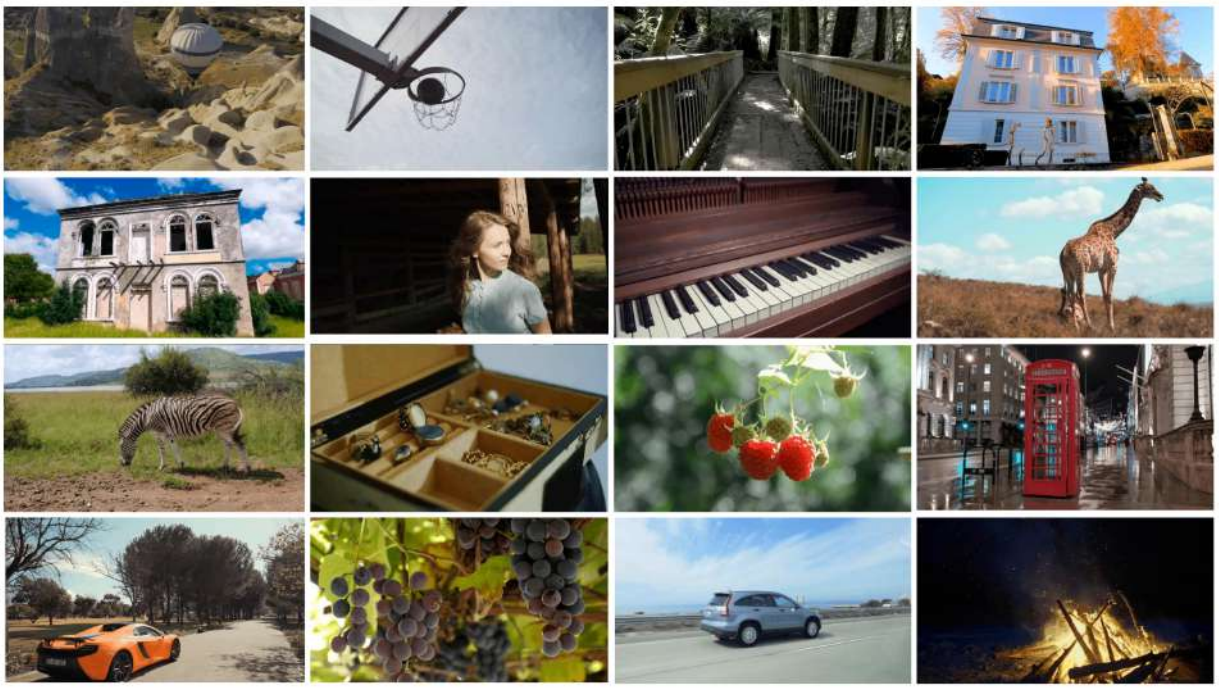

Figure 5.11: **Stimulus set.** Example stimuli used in the experiment. The shared-category items consisted of two exemplars each from four categories: cars, houses, animals, and fruits. All original images were sourced from the free stock image database Pexels. One stimulus depicting a human was replaced with a schematic human icon to comply with privacy and identifiability restrictions.

Decoding performance

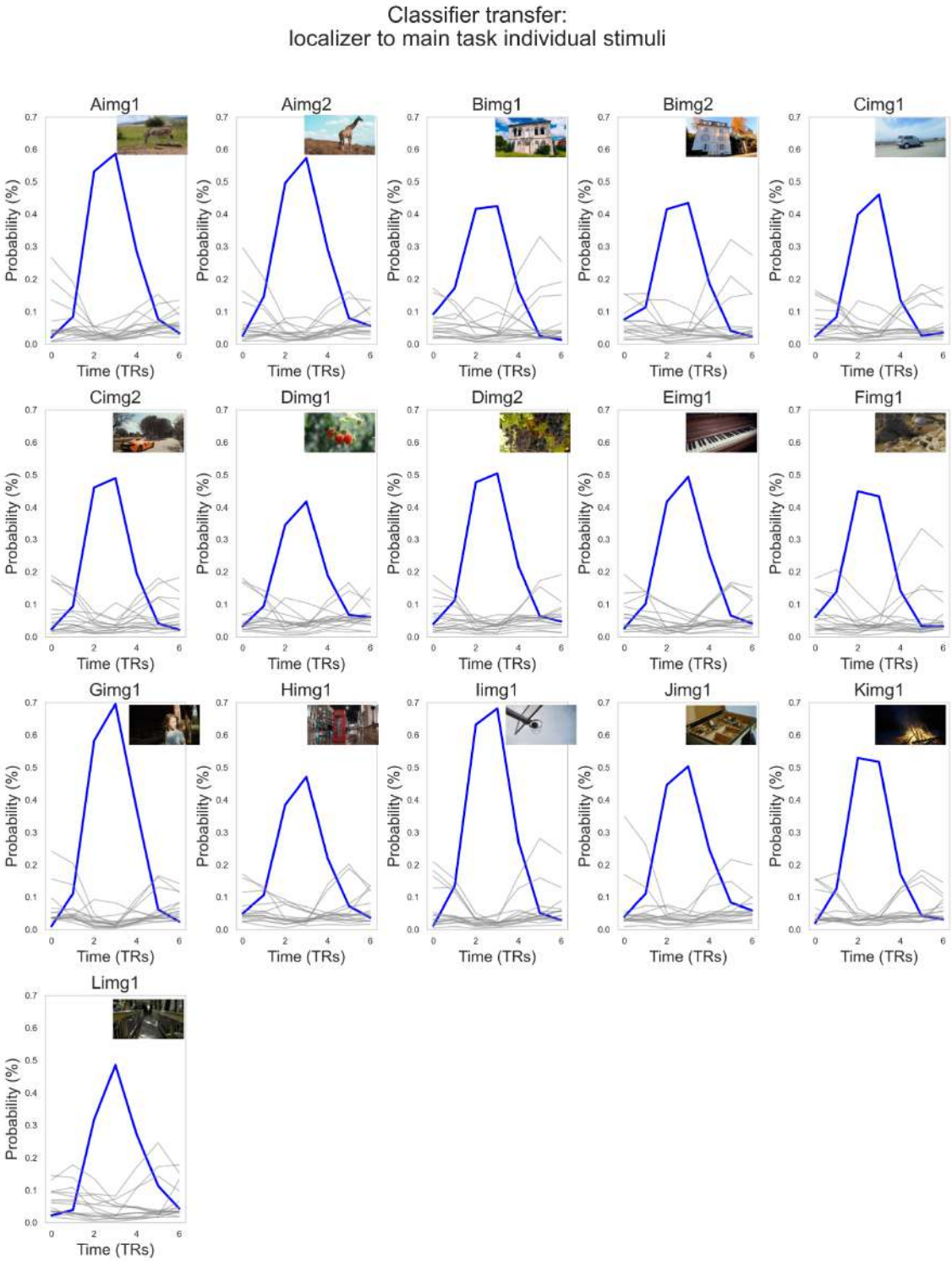

Figure 5.12: **Classifier transfer from localizer to main task.** Time-resolved classifier outputs (in TRs) for each of the 16 stimuli, decoded from main task data using the localizer-trained classifier. For each stimulus, the corresponding class (blue) exhibits a strong, stimulus-locked probability increase, while nonmatching classes (grey) remain low. This pattern demonstrates robust stimulus-specific decoding and successful transfer of the classifier from the localizer scans to the main task.

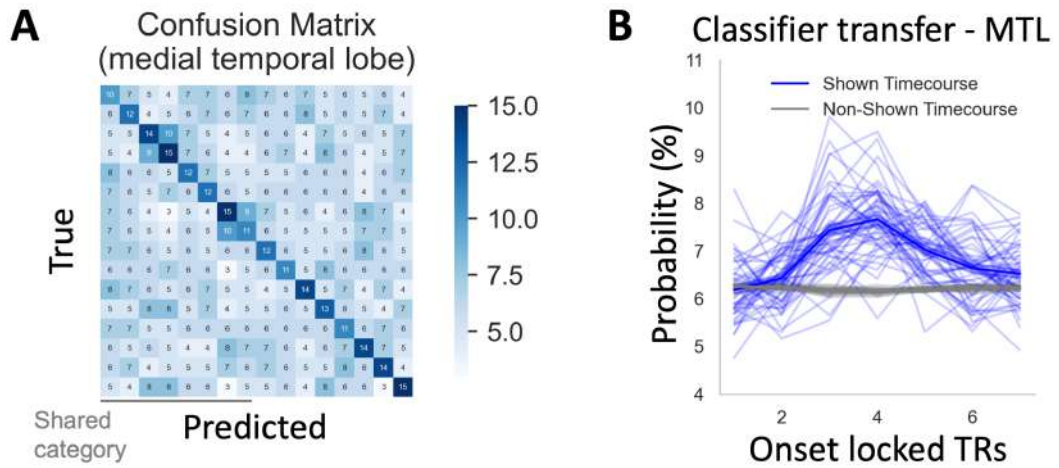

Figure 5.13: **Decoding for medial temporal lobe.** (A) Confusion matrix for all 16 stimuli decoded within the medial temporal lobe ROI. A clear diagonal indicates successful stimulus-level decoding despite overall lower decoding accuracy in this region. Off-diagonal structure shows some degree of category-level confusion, yet the dominant diagonal pattern confirms that stimulus-specific information is still reliably recovered. (B) Classifier transfer to the main task: classifier probabilities show a peak following stimulus presentation for shown items (blue) but not non-shown items (gray).

Replay results day 2 visual cortex before structural change

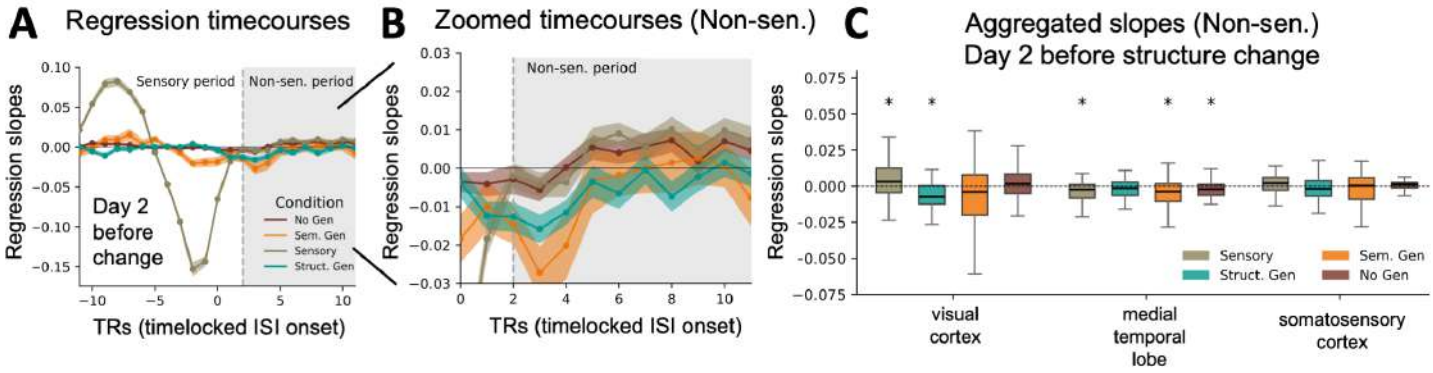

Figure 5.14: **Replay evidence in visual cortex on Day 2 before the structural change.** (**A–B**) Regression timecourses timelocked to ISI onset show replay dynamics during the sensory and non-sensory periods. The pattern closely resembles Day 1 (Fig. 5.5 in the main text): during the sensory period, activity tracks the currently observed stimulus, and during the subsequent non-sensory period, the reward-linked sequences are reactivated. (**C**) Aggregated regression slopes from the non-sensory period reproduce the main characteristics observed on Day 1. Statistical testing here is based only on the first two blocks of Day 2 (i.e., before the hidden structure change) and is therefore less powered, but the overall replay profile remains consistent with the main-text results.

Chapter 6

Decoding sleep's role in memory transformation: Sleep-dependent reorganization of sequential replay during retrieval

Fabian M. Renz, Marit Petzka, Elsa Kolbe, Christian F. Doeller, Nicolas W. Schuck

In preparation.

6.1 Introduction

Sleep actively transforms memory. An early and influential account held that sleep benefits memory primarily through passive protection from retroactive interference. Wakefulness continually exposes recently encoded traces to new, partially overlapping experiences that can overwrite or disrupt them, whereas sleep reduces this exposure and thereby preserves memories that would otherwise decay through interference [Jenkins and Dallenbach, 1924, Diekelmann and Born, 2010]. On this view, the mnemonic benefit of sleep follows from the absence of interfering input, rather than from any process sleep actively performs. Behavioural evidence has since shifted the field toward a more active view: memories consolidated across sleep are not merely preserved but rendered more resistant to subsequent interference [Ellenbogen et al., 2006, 2009], and lesion and imaging work indicates that their dependence on the hippocampus declines over time as neocortical representations become sufficient to support retrieval [Frankland and Bontempi, 2005]. Together these findings suggest that sleep promotes the gradual redistribution of memories from hippocampal to neocortical networks, a process during which their representational content is reorganized [Diekelmann and Born, 2010, Rasch and Born, 2013, Klinzing et al., 2019, Lutz et al., 2026]. The complementary learning systems (CLS)

framework provides the theoretical foundation for this view. Connectionist networks trained sequentially on overlapping material catastrophically overwrite earlier learning [McCloskey and Cohen, 1989, French, 1999], and McClelland, McNaughton, and O'Reilly turned this failure into a functional argument: a single system cannot simultaneously support rapid encoding of arbitrary new experiences and slow extraction of statistical structure across experiences without one disrupting the other [McClelland et al., 1995]. The brain is proposed to resolve this tension through two specialised systems: 1) a fast, pattern-separated encoding in the hippocampus and 2) a slow, statistical learning in the neocortex, coupled by offline reactivation that interleaves new information with existing cortical knowledge [McClelland et al., 1995, Kumaran et al., 2016]. The transfer of information between these systems should therefore not merely relocate memories but change what they represent, shifting from detailed episodic traces toward more abstract, schema-integrated formats [Frankland and Bontempi, 2005, Brodt et al., 2023]. Neural replay during sleep is the leading candidate mechanism for this transfer: hippocampal sharp-wave ripple-associated reactivation is causally required for the consolidation of hippocampus-dependent memory [Girardeau et al., 2009], and its coordination with neocortical slow oscillations and thalamocortical sleep spindles provides the temporal scaffolding for hippocampal-to-cortical communication [Born et al., 2006, Klinzing et al., 2019].

Substantial evidence supports the anatomical prediction that retrieval shifts from hippocampal to cortical dependence over time [Gais et al., 2007, Takashima et al., 2009, Brodt et al., 2018, Ko et al., 2025], and multivariate neuroimaging studies have begun to show that consolidated memories come to reflect shared structure across related experiences [Tompary and Davachi, 2017, 2024]. However, this work has focused almost exclusively on simple associative memories such as word pairs, object-location bindings, or category exemplars. How consolidation transforms more complex, hierarchically structured memories remains largely unknown.

This gap is significant because the episodes we encode in daily life are rarely isolated associations. They are embedded within nested contexts that organize experience at multiple scales. A typical workday, for example, unfolds within a particular city and building, comprises a series of distinct events (a morning meeting, a conversation in the hallway, lunch at a nearby café), each with its own ordered sequence of people, objects, and actions. Memory for such episodes therefore requires representing information at multiple levels simultaneously: the specific details of individual events, the temporal order in which they occurred, and the higher-order contextual structure that binds them together. Hierarchical organization of this kind is not a special case but a pervasive feature of naturalistic experience [Zacks and Swallow, 2007].

This hierarchical structure is of particular interest for theories of consolidation because it places differential demands on the two memory systems posited by the CLS framework. Encoding fine-grained sequential detail, such as the specific order of items within an event, requires the kind of pattern-separated, high-fidelity representations attributed to the hippocam-

pus. Extracting higher-order regularities, such as which events belong to the same context or which categorical structure recurs across different episodes, engages the kind of slow statistical learning attributed to the neocortex. A hierarchical sequence memory task therefore provides a particularly informative test of the systems consolidation hypothesis, because it requires both fine-grained sequential encoding and higher-order contextual organization.

A further open question concerns the neural signatures of sequential memory retrieval and how they change across consolidation. Several recent studies have demonstrated that fMRI can detect temporally ordered reactivation of task states during rest and retrieval [Schuck and Niv, 2019, Wittkuhn and Schuck, 2021, Wittkuhn et al., 2025]. The preceding chapters of this thesis used these methods to show that replay during wakefulness operates over inferred latent structure, with replay strength tracking non-local value updating on a trial-by-trial basis (Chapter 2). However, whether the locus and character of such sequential reactivation change across a night of sleep, as the systems consolidation framework would predict, has not been tested. If replay during sleep drives representational reorganization, then the neural signatures of retrieval itself should differ before and after consolidation, potentially shifting from hippocampal to cortical dependence.

The present study addresses these questions by combining overnight fMRI with multivariate decoding in a hierarchical sequence memory task. Participants learned ordered sequences of objects embedded within scenes and organized into overarching contexts, creating a three-level memory structure. Retrieval was assessed in the scanner both before and after a night of sleep. Multivariate classifiers were trained on stimulus-driven neural activity from independent localizer sessions, in which participants viewed task images paired with unrelated words and subsequently recalled images from word cues. These classifiers were then applied to the main task retrieval sessions to decode stimulus-specific reactivation patterns across cortical and medial temporal regions of interest.

The key methodological contribution of this study is the application of the sequential slope ordering analysis (SODA) metric [Wittkuhn and Schuck, 2021, Renz et al., 2026] to the study of sequential retrieval. The standard SODA approach quantifies sequentiality by regressing sequence position onto decoded probabilities at each time point; taking the absolute value of the resulting slope captures sequential structure regardless of replay direction. This absolute SODA metric is the primary measure used throughout this chapter. As a complementary, phase-invariant measure, we additionally fit a sinusoidal model to each trial's SODA time course, yielding a sinusoidal amplitude (SinAmp) that captures retrieval strength regardless of temporal onset variability across trials (see Methods for details). We report both metrics throughout and examine their convergence.

This design allows a direct comparison of the locus and strength of sequential memory reactivation before and after overnight consolidation. We test two predictions derived from the systems consolidation framework. First, sequential reactivation, the temporally ordered reinstatement of encoded sequence elements assessed through multivariate decoding, should

be detectable during retrieval trials and should predict retrieval success. Second, the regional profile of sequential reactivation should change across sleep, with post-sleep retrieval showing relatively stronger cortical contributions compared to pre-sleep retrieval.

6.2 Results

The results presented in this chapter are part of a larger collaborative project with Dr. Marit Petzka and Elsa Kolbe, conducted between the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig and the University of Hamburg. This chapter reports the fMRI retrieval analyses that constitute my primary contribution to the project, focusing on the pre- versus post-sleep comparison of sequential reactivation. Additional analyses, including computational modelling of representational change, will be reported in the forthcoming publication.

6.2.1 Experimental design

To investigate how sleep transforms memory representations and whether neural replay tracks these changes, we designed a two-day fMRI experiment centered on a hierarchical sequence memory task (Fig. 6.1).

Participants learned ordered sequences of everyday objects organized into a three-level hierarchical structure (Fig. 6.1B). At the highest level, two overarching *contexts* (day vs. night, indicated by background color) each contained five distinct *scenes* (e.g., autumn forest, waterfall, desert). Within each scene, participants learned an ordered sequence of five unique objects (e.g., dog, cardigan, piano, basketball, table), yielding 50 stimuli in total (2 contexts $\times$ 5 scenes $\times$ 5 objects). Two distinct category sequences were used to assign objects to positions within scenes: sequence A–E (e.g., animal, clothing, instrument, sports equipment, furniture) and sequence F–J (e.g., a different set of five semantic categories). Critically, both category sequences appeared in both contexts: within each context, three scenes followed category sequence A–E and two scenes followed sequence F–J (or vice versa), ensuring that category structure was not confounded with context membership (see letters in Fig. 6.1B).

The experiment spanned two days with multiple MRI sessions (Fig. 6.1A). Two independent *localizer sessions* were conducted before the main task, introducing all 60 task stimuli (50 objects and 10 scenes) in a word–image association paradigm used to train multivariate classifiers (described below). On the core experimental day, participants completed the main task *encoding session* in the evening. During encoding, participants first studied the ordered sequences of scenes within each context (presented three times) and then the ordered sequences of objects within each scene (presented twice).

Following encoding, participants completed a *pre-sleep retrieval session* (Ret_1), testing memory for both scene order and object order (Fig. 6.1C). On scene recall trials, participants were shown the context together with a position cue and asked to recall which scene appeared

at that position. On object recall trials, participants were shown the context and scene along with a position cue and asked to recall the specific object. Crucially, on each trial participants first indicated their retrieval success via button press and then provided a verbal free-recall response, without receiving any feedback or seeing the correct answer. This prevented reliance on recognition-based strategies, ensured access to genuine sequence memory, and avoided any re-encoding of the correct solution during retrieval.

After the pre-sleep retrieval session, participants slept overnight, first for up to three hours in the scanner and then outside for the remainder of the night. Analysis of the sleep recordings is part of a separate project and not included here. The next morning, participants completed two full *post-sleep retrieval sessions* (Ret_2), each identical in structure to the pre-sleep session, providing two complete retrievals of all items for additional statistical power in the MRI analyses. This within-participant design, with matched retrieval sessions flanking a night of sleep, enables a direct comparison of the neural signatures of memory retrieval before and after sleep-dependent consolidation.

In the localizer sessions, participants studied word–image pairs for all 60 task stimuli (50 objects and 10 scenes) and subsequently recalled images from word cues alone. Logistic regression classifiers were trained on the neural patterns evoked during both the localizer encoding blocks (when participants directly viewed the images) and the localizer retrieval blocks (when participants recalled images from word cues), collapsing across individual object identities to create 10 object category labels (e.g., dogs, cats, and other animals were summarized as "animals") and 10 scene labels. The motivation for including retrieval trials in classifier training was to capture neural patterns more closely aligned with the top-down process of memory retrieval, as it occurs when sequences are embedded in the main task. The resulting classifiers were applied, without retraining, to the encoding and retrieval sessions of the main task, producing category-specific probability time courses at each TR (see Methods for details of the classifier pipeline and voxel selection procedures).

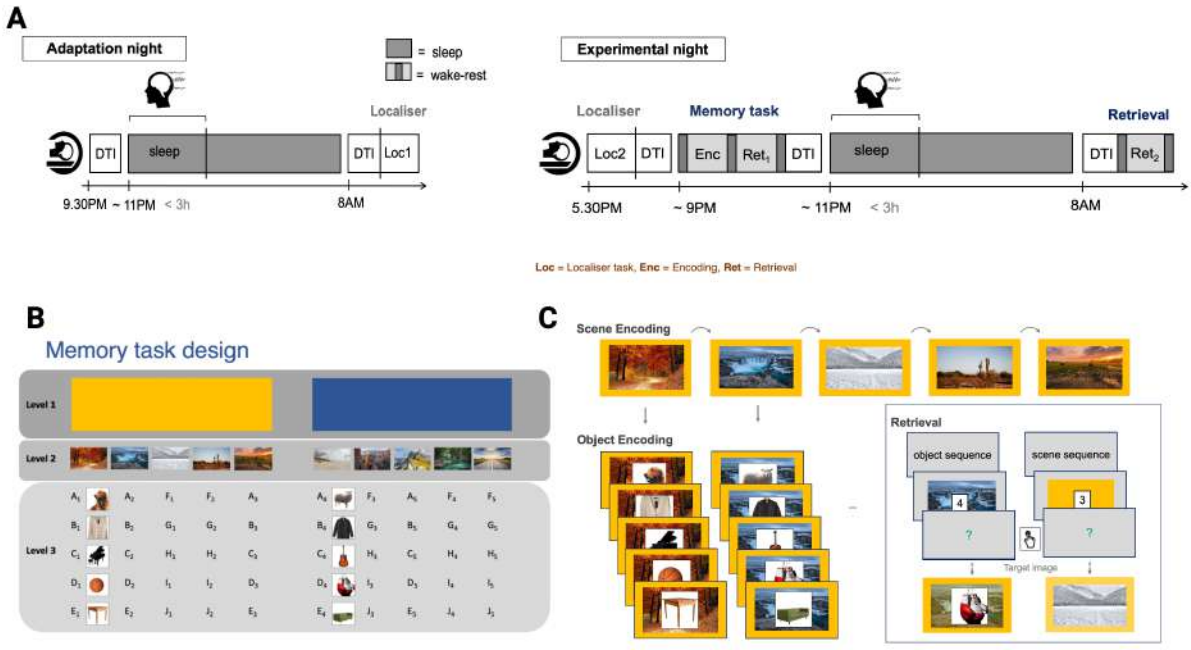

Figure 6.1: **Experimental design and hierarchical task structure.** **(A)** Timeline. The figure shows the complete data collection protocol. Two localizer sessions (Loc1, Loc2) were conducted on separate days. On the experimental day: encoding (Enc), pre-sleep retrieval (Ret_1; evening), overnight sleep, and two post-sleep retrieval sessions (Ret_2; morning). Note that the collected structural MRI scans and nighttime sleep measurements (EEG during in-scanner sleep) are part of separate projects and are not analyzed in this chapter; the present work focuses exclusively on the pre- versus post-sleep retrieval comparison. **(B)** Hierarchical task structure. Participants learned sequences organized across three nested levels: two contexts (day/night, indicated by yellow/blue background), each containing five ordered scenes (Level 2), each containing an ordered sequence of five objects (Level 3; 50 stimuli total). Two distinct category sequences (A–E and F–J) assigned semantic categories to object positions; both category sequences appeared in both day and night contexts (three scenes of one sequence and two of the other per context). **(C)** Encoding and retrieval design. During encoding, participants first studied ordered scene sequences (top), then studied ordered object sequences embedded within each scene (bottom left). During retrieval (bottom right), object recall trials presented the context, scene, and a position cue; scene recall trials presented the context and a position cue. Participants indicated retrieval success via button press and then gave a verbal free-recall response.

6.2.2 Behavioral performance

Participants ($N = 39$) performed the hierarchical memory task with high accuracy (Fig. 6.2). Scene retrieval was near ceiling, with a mean accuracy of 97% across all retrieval sessions and sequence positions. Object retrieval was lower but still well above chance, with a mean accuracy of 81% (Fig. 6.2A). This difference is consistent with the hierarchical task design, in which scenes served as higher-order contextual anchors that were encoded more frequently (three vs. two repetitions) and probed with simpler cues (context + position) compared to objects (context + scene + position).

Serial position analyses revealed a monotonic decline in object retrieval accuracy from position 1 ($\sim$89%) to position 5 ($\sim$75%), consistent with a primacy-dominated gradient in which early sequence items benefit from additional rehearsal and reduced proactive interference (Fig. 6.2B). Scene retrieval accuracy was near ceiling across all five positions ($\geq$95%), showing no reliable serial position effect. Reaction times mirrored this pattern: object retrieval was slower overall than scene retrieval ($\sim$1.9 s vs. $\sim$1.4 s) and showed an inverted-U serial position curve peaking around position 4, consistent with slower and more effortful retrieval for middle and late sequence positions (Fig. 6.2C,D).

Critically, retrieval accuracy was stable across the pre-sleep evening session (Ret_1) and the two post-sleep morning sessions (Ret_2), for both scenes and objects. Scene accuracy remained at ceiling ($\sim$94–98%) and object accuracy remained at $\sim$79–82% across all three sessions. A linear mixed-effects model (with session as fixed effect and random intercepts for participants) revealed a small but statistically significant increase in overall retrieval rate from evening to morning (Morning 1 vs. Evening: $b = 2.38$, $z = 3.50$, $p < .001$; Morning 2 vs. Evening: $b = 2.74$, $z = 4.02$, $p < .001$), though the magnitude of this improvement was modest ($\sim$2–3 percentage points). Reaction times decreased from the evening to the morning sessions for both memory types (Morning 1 vs. Evening: $b = -0.52$ s, $z = -9.72$, $p < .001$; Morning 2 vs. Evening: $b = -0.67$ s, $z = -12.47$, $p < .001$). This relative behavioral stability in accuracy across sessions is important for interpreting the neural data: any changes in retrieval-related brain activity between pre- and post-sleep sessions are unlikely to reflect large differences in memory strength, but instead suggest reorganization in how the same memories are neurally represented and retrieved.

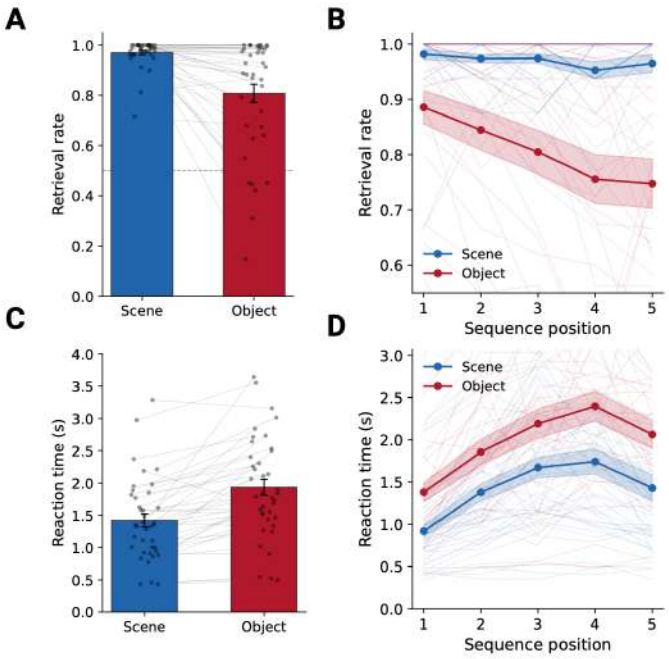

Figure 6.2: **Behavioral performance in the hierarchical memory task.** **(A)** Mean retrieval accuracy for objects (red) and scenes (blue). $N = 39$. **(B)** Retrieval accuracy by sequence position (1–5) for scenes and objects. Scene accuracy is near ceiling across positions; object accuracy declines monotonically from position 1 to position 5. **(C)** Mean reaction times (RT) for object and scene retrieval. **(D)** RT by sequence position, showing an inverted-U pattern for object retrieval. Error bars denote ± 1 SEM; gray lines connect individual participants. Session-level accuracy and reaction time data are reported in the main text.

The high overall accuracy, combined with the graded difficulty across hierarchical levels and serial positions, provides a rich behavioral landscape for relating neural measures to memory performance. At the same time, the near-ceiling scene accuracy limits the number of forgotten scene trials available for subsequent memory contrasts, a constraint we address in the multivariate analyses below by focusing primarily on object retrieval.

Of the 46 participants recruited, several were excluded at various stages due to technical difficulties, including incomplete MRI sessions and missing behavioral data. To maximize statistical power, each analysis includes the largest number of participants with complete data for that specific pipeline: $N = 39$ for behavioral analyses, $N = 35$ for classifier cross-validation, $N = 34$ for retrieval decoding, and encoding transfer analysis.

6.2.3 Classifier training and cross-session generalization

We next turned to multivariate pattern analysis to assess the fidelity of stimulus-specific neural representations and their reactivation during retrieval. Logistic regression classifiers were trained on both the encoding and retrieval blocks of the localizer sessions to decode stimulus categories (10 scenes, 10 object categories) from fMRI activity patterns in each ROI (see Methods). By including retrieval trials, the classifiers were trained on neural patterns that more

closely reflect the top-down memory retrieval process engaged during the main task. Classifier performance was evaluated using stratified leave-one-run-out cross-validation ($N = 35$).

Classification accuracy was well above the 10% chance level in all three ROIs for both stimulus types (Fig. 6.3A). For scenes, visual cortex achieved the highest accuracy (30.4%; $t(34) = 12.91$, $p < 0.001$), followed by parietal cortex (20.2%; $t(34) = 8.35$, $p < 0.001$) and MTL (12.4%; $t(34) = 7.18$, $p < 0.001$). For objects, the pattern was similar: visual cortex 29.3% ($t(34) = 14.21$, $p < 0.001$), parietal cortex 19.6% ($t(34) = 9.97$, $p < 0.001$), and MTL 13.4% ($t(34) = 11.25$, $p < 0.001$). The gradient from visual cortex to parietal to MTL is consistent with the expectation that encoding-trained classifiers capture stimulus-driven signals most strongly in early visual regions, with weaker but reliable decoding in higher-order areas.

To verify that the localizer-trained classifiers generalized to the main task, we applied them to the encoding session and examined whether decoded probabilities for the presented stimulus peaked at the expected time (Fig. 6.3B). For both scenes and objects, the classifier probability for the target class rose following stimulus onset, peaking approximately 6 s later consistent with the hemodynamic delay, and decayed afterward. Peak decoding accuracy was significantly above the 10% chance level in all three ROIs for both stimulus types ($N = 33$): visual cortex achieved 23.9% for objects ($t(32) = 14.25$, $p < 0.001$, $d = 2.48$) and 19.1% for scenes ($t(32) = 14.32$, $p < 0.001$, $d = 2.49$); parietal cortex achieved 15.3% for objects ($t(32) = 9.72$, $p < 0.001$, $d = 1.69$) and 12.7% for scenes ($t(32) = 10.22$, $p < 0.001$, $d = 1.78$); and MTL achieved 11.6% for objects ($t(32) = 5.94$, $p < 0.001$, $d = 1.03$) and 10.6% for scenes ($t(32) = 3.33$, $p = 0.002$, $d = 0.58$). This confirms that the classifiers successfully tracked stimulus-specific neural patterns during the main task, validating their use in the subsequent retrieval analyses.

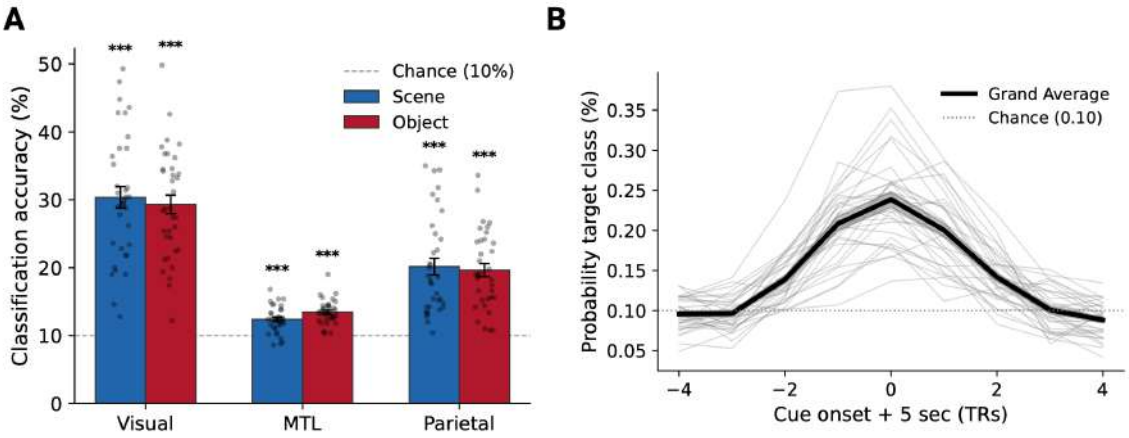

Figure 6.3: **Classifier training and cross-session generalization. (A)** Cross-validated classification accuracy for scenes and objects in visual cortex, parietal cortex, and MTL. All ROIs significantly exceed chance (10%; all $p < 0.001$). $N = 35$. **(B)** Time course (visual ROI) of decoded target-class probability during encoding, locked to stimulus onset (+5 s HRF delay). Probabilities peak following cue presentation, confirming cross-session generalization.

6.2.4 Stimulus-specific reinstatement during retrieval

Having established that the classifiers generalized to stimulus-driven activity during encoding, we next asked whether they could detect stimulus-specific reinstatement during retrieval, when participants recalled sequence items from memory without seeing the corresponding images. For each retrieval trial, we computed the target-class probability at each TR and averaged this timecourse across trials, then extracted the peak of the resulting mean timecourse as a measure of reinstatement strength (analogous to the encoding transfer analysis above).

Collapsing across sessions, peak target-class probability was significantly above chance (10%) in all three ROIs ($N = 34$): visual cortex (14.9%; $t(33) = 13.39$, $p < 0.001$), parietal cortex (13.5%; $t(33) = 15.47$, $p < 0.001$), and MTL (12.7%; $t(33) = 23.58$, $p < 0.001$).

These results confirm that the classifiers successfully detected content-specific memory reinstatement during retrieval: even in the absence of visual input, neural patterns carried information about which item participants were recalling. This reinstatement signal provides the foundation for the sequential analyses below, which test whether the decoded category probabilities further respect the temporal ordering of the encoded sequences.

6.2.5 Sequential reactivation during retrieval

Having established that classifiers generalized to retrieval sessions and detected stimulus-specific reinstatement, we next asked whether memory retrieval reinstated the temporal order of encoded sequences. We quantified sequential ordering using the sequential slope ordering analysis (SODA) [Wittkuhn and Schuck, 2021, Wittkuhn et al., 2025, Renz et al., 2026], which regresses sequence position onto decoded probabilities at each time point, yielding a slope that indexes forward or backward sequential reactivation (see Methods). By taking the absolute value of the slope, we capture sequentiality regardless of replay direction. The primary comparison contrasts the absolute SODA slope for the retrieved sequence against that of a control sequence. For scene retrieval trials, the control was the scene sequence from the other context (e.g., if the day scene order was probed, the night scene order served as control). For object retrieval trials, the control was the object sequence drawn from the other category sequence (e.g., if the probed scene contained objects following category sequence A–E, the control was the corresponding F–J sequence, or vice versa).

Because the temporal onset of sequential retrieval varies across trials, averaging raw SODA slopes can lead to phase cancellation. If retrieval of a five-item sequence unfolds as a temporally ordered sweep through decoded probabilities, each trial's SODA time course should approximate a sinusoidal oscillation whose phase depends on when retrieval begins relative to the cue. Averaging within or across trials with different phases can cancel the signal entirely, producing a flat mean even when individual trials contain robust sequentiality. To address this, we computed the absolute SODA, avoiding positive and negative phases cancelling one another, and the sinusoidal amplitude (SinAmp): for each trial, a sinusoidal model was fitted to

the SODA time course, and the fitted amplitude captures retrieval strength regardless of phase alignment across trials (see Methods).

Sequential reactivation was detectable in all three ROIs when collapsing across sessions (Fig. 6.4A). For the absolute SODA metric, the replay effect (retrieved − control) was significantly greater than zero in visual cortex ($t(33) = 2.70$, $p = 0.011$), medial temporal lobe ($t(33) = 2.88$, $p = 0.007$), and parietal cortex ($t(33) = 2.12$, $p = 0.042$). The SinAmp metric confirmed this pattern, with significant effects in all three ROIs (visual cortex: $t(33) = 2.42$, $p = 0.021$; MTL: $t(33) = 2.89$, $p = 0.007$; parietal: $t(33) = 3.14$, $p = 0.004$). All six tests survived FDR correction across ROIs and metrics (all $p_{\text{FDR}} < 0.043$), confirming that sequential reactivation during retrieval is a robust phenomenon.

The session-level breakdown revealed a dissociation across regions (Fig. 6.4B). During pre-sleep retrieval, significant sequential reactivation was observed only in visual cortex (SODA: $t(32) = 2.47$, $p = 0.019$, $p_{\text{FDR}} = 0.046$), while MTL ($p = 0.307$) and parietal cortex ($p = 0.808$) did not reach significance. During post-sleep retrieval, the pattern shifted: MTL showed robust sequentiality ($t(32) = 2.76$, $p = 0.009$, $p_{\text{FDR}} = 0.034$) and parietal cortex emerged as significant ($t(32) = 2.69$, $p = 0.011$, $p_{\text{FDR}} = 0.034$), while visual cortex did not reach significance ($t(32) = 1.86$, $p = 0.072$). FDR correction was applied across 12 tests (3 ROIs $\times$ 2 sessions $\times$ 2 metrics). This regional shift is partially consistent with the systems consolidation prediction: the emergence of parietal replay after sleep aligns directly with the expected redistribution of memory traces toward cortical regions. The MTL result is more nuanced: the CLS framework would predict a *decrease* in hippocampal dependence after consolidation, yet MTL replay emerged post-sleep rather than declining. This may reflect the ongoing role of the hippocampus in coordinating retrieval even after partial cortical takeover, or it may be inflated by the fact that post-sleep had twice as many trials as pre-sleep, a power asymmetry we address in a dedicated control analysis below.

The replay effect did not differ significantly between object and scene trials (LME: all $p > 0.14$), but the 50 unique objects across 10 sequences carry the bulk of behavioral variance, whereas the 5 scenes are retrieved at ceiling ($\sim$97%). We therefore focus all subsequent analyses on object trials, where individual differences in memory performance are most informative.

As a validation, we repeated the detection analysis using shuffled sequence position labels, which destroy the temporal ordering while preserving item-level information. Many shuffled comparisons remained significant (11/18 tests at $p < 0.05$), indicating that the absolute SODA metric captures a composite signal reflecting both sequential ordering and general reinstatement strength. The implications of this finding for the interpretation of the replay effects are discussed in detail in the Discussion.

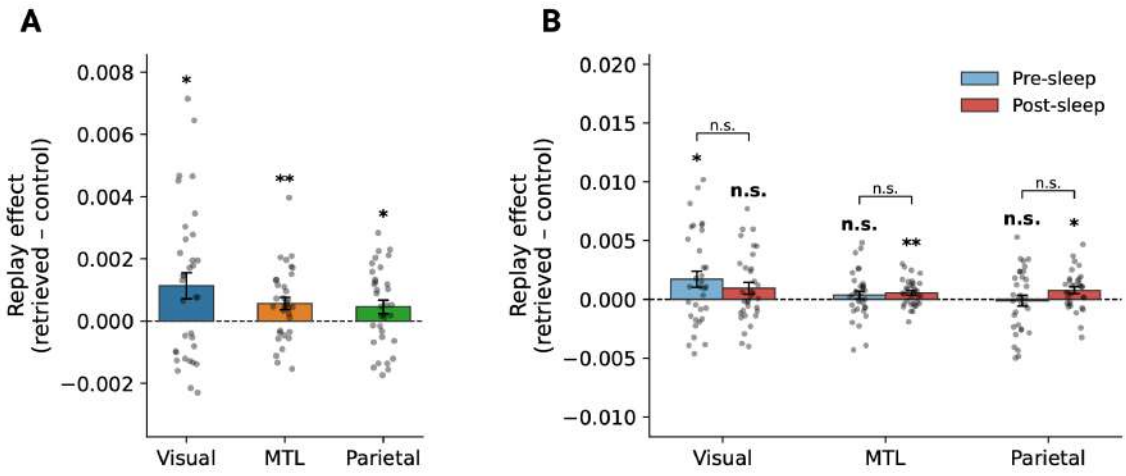

Figure 6.4: **Sequential reactivation during retrieval.** **(A)** Absolute SODA replay effect (retrieved − control) collapsed across sessions for visual cortex, MTL, and parietal cortex. All three ROIs show significant replay. The SinAmp metric confirmed this pattern (see main text). **(B)** Session breakdown. The metric shown is the absolute SODA slope (retrieved − control). Pre-sleep: visual cortex significant ($p = 0.019$, $p_{\text{FDR}} = 0.046$); MTL and parietal not significant. Post-sleep: MTL ($p = 0.009$, $p_{\text{FDR}} = 0.034$) and parietal ($p = 0.011$, $p_{\text{FDR}} = 0.034$) significant; visual cortex not significant ($p = 0.072$).

6.2.6 Sleep-dependent changes in sequential reactivation

The preceding section established that sequential reactivation during retrieval is detectable across brain regions, but hinted at a regional shift between sessions: visual cortex dominated before sleep, while MTL and parietal cortex emerged after sleep. The central prediction of the systems consolidation framework is that sleep-dependent replay drives the redistribution of memory representations from hippocampus to cortex. If this is the case, the strength of sequential reactivation should increase from pre-sleep to post-sleep in cortical regions. We tested this directly by comparing replay strength before and after sleep.

Because participants completed two post-sleep retrieval blocks but only one pre-sleep block, we report two complementary analyses: a full-data analysis using all available trials and a trial-matched analysis restricted to the first post-sleep block (runs 1–3), which equates trial counts across sessions.

For scene trials, per-ROI session comparisons (LME on replay difference) suggested an increase in MTL replay from pre-sleep to post-sleep ($z = 2.25$, $p = 0.025$ uncorrected; $p_{\text{FDR}} = 0.075$; Fig. 6.5A), though this effect did not survive correction for multiple comparisons across the three ROIs. Visual cortex ($p = 0.214$) and parietal cortex ($p = 0.633$) showed no change. Given the near-ceiling behavioral accuracy for scenes ($\sim$97%), we do not pursue this effect further here, but note that the direction is consistent with the MTL pattern observed for objects (see below). For object trials, we fitted a LME predicting the replay metric from retrieval condition (retrieved vs. control), session (pre-sleep vs. post-sleep), their interaction, and ROI

to test regional specificity. In the full-data analysis, the three-way retrieval × session × ROI interaction was significant for parietal cortex relative to MTL (SODA: $z = 2.13$, $p = 0.033$) but not for visual cortex ($z = 0.02$, $p = 0.981$), indicating that the session effect was specific to parietal cortex. The trial-matched analysis confirmed and strengthened this result: the parietal retrieval × session interaction was significant for both metrics after FDR correction across the three ROIs (absolute SODA: $z = 3.04$, $p = 0.002$, $p_{\text{FDR}} = 0.014$; SinAmp: $z = 2.76$, $p = 0.006$, $p_{\text{FDR}} = 0.018$; Fig. 6.5B), indicating that the gap between retrieved and control sequences specifically widened after sleep. Visual cortex and MTL showed no such interaction in either analysis (all $p > 0.44$).

Fig. 6.5C illustrates the pattern of this interaction for parietal cortex: object replay was near zero before sleep and increased after sleep, with a consistent pattern across both the absolute SODA and SinAmp metrics.

The parietal replay signal was concentrated in the first post-sleep retrieval block and subsequently decayed: comparing early (runs 1–3) versus late (runs 4–6) post-sleep runs revealed a significant decline in parietal cortex (SinAmp: $t(32) = -2.85$, $p = 0.008$, $d = -0.50$), with no decay in visual cortex or MTL (all $p > 0.18$).

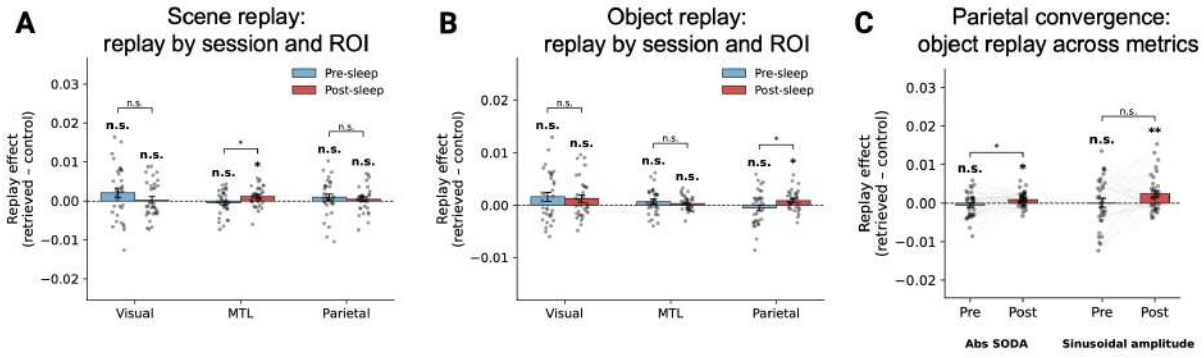

Figure 6.5: **Sleep-dependent emergence of parietal sequential reactivation.** **(A)** Scene trials: absolute SODA replay difference (retrieved − control) for pre-sleep and post-sleep sessions across ROIs. Brackets show per-ROI session comparisons (LME). MTL shows an increase post-sleep ($p = 0.025$ uncorrected; $p_{\text{FDR}} = 0.075$); visual cortex and parietal cortex show no change. **(B)** Object trials: absolute SODA replay difference (retrieved − control) for pre-sleep and post-sleep sessions across ROIs. Parietal cortex shows a region-specific session effect (retrieval × session × ROI: $p = 0.033$; trial-matched retrieval × session interaction, SODA: $p = 0.002$; SinAmp: $p = 0.006$). Visual cortex and MTL show no session effect. **(C)** Parietal cortex only: pre- versus post-sleep replay for objects, shown separately for absolute SODA (left) and SinAmp (right). Stars above individual bars denote one-sample tests against zero; brackets show session comparisons (LME).

6.2.7 Brain–behavior associations

If sequential reactivation during retrieval supports memory, participants with stronger replay should show better behavioral performance. We computed subject-level Pearson correlations

between replay metrics and mean retrieval accuracy for object trials.

The replay difference (retrieved − control) correlated positively with retrieval accuracy in visual cortex ($r = 0.370$, $p = 0.031$, $p_{\text{FDR}} = 0.047$; Fig. 6.6A) and parietal cortex ($r = 0.449$, $p = 0.008$, $p_{\text{FDR}} = 0.024$; Fig. 6.6B), such that participants with stronger replay in these regions retrieved more objects correctly. Both associations survived FDR correction across the three ROIs. The parietal association was already present in the pre-sleep session ($r = 0.389$, $p = 0.025$). In MTL, the replay difference showed a trend toward predicting accuracy ($r = 0.331$, $p = 0.060$; Fig. 6.6C); notably, using the raw retrieved signal (without subtracting control) yielded a significant association ($r = 0.372$, $p = 0.030$; collapsed across sessions), which strengthened post-sleep ($r = 0.416$, $p = 0.016$). The SinAmp metric showed a convergent pattern in MTL post-sleep ($r = 0.359$, $p = 0.040$). All subsidiary correlations survived FDR correction across the seven reported brain–behavior tests (all $p_{\text{FDR}} < 0.047$).

Importantly, replay strength did not differ between correctly and incorrectly remembered trials at the trial level (LME: all $p > 0.14$), despite the between-subject associations documented above. This dissociation suggests that replay reflects an individual difference: participants who engage in stronger sequential reactivation tend to remember better overall. The null trial-level accuracy effect may also reflect limited power given the strong imbalance between correct ($\sim$85%) and incorrect ($\sim$15%) trials.

As a final validation, we examined whether the sinusoidal model fit quality (R^2) itself differed between retrieved and control trials. If replay produces a genuinely periodic SODA time course, the sinusoidal model should fit better for retrieved sequences. This was confirmed: the retrieved R^2 was significantly higher than the control R^2 in visual cortex ($t(33) = 2.28$, $p = 0.029$), with parietal cortex showing a trend ($t(33) = 2.01$, $p = 0.053$). Furthermore, the two replay metrics were strongly correlated across participants in all three ROIs (Fig. 6.7), confirming that the absolute SODA and SinAmp measures capture converging aspects of sequential reactivation.

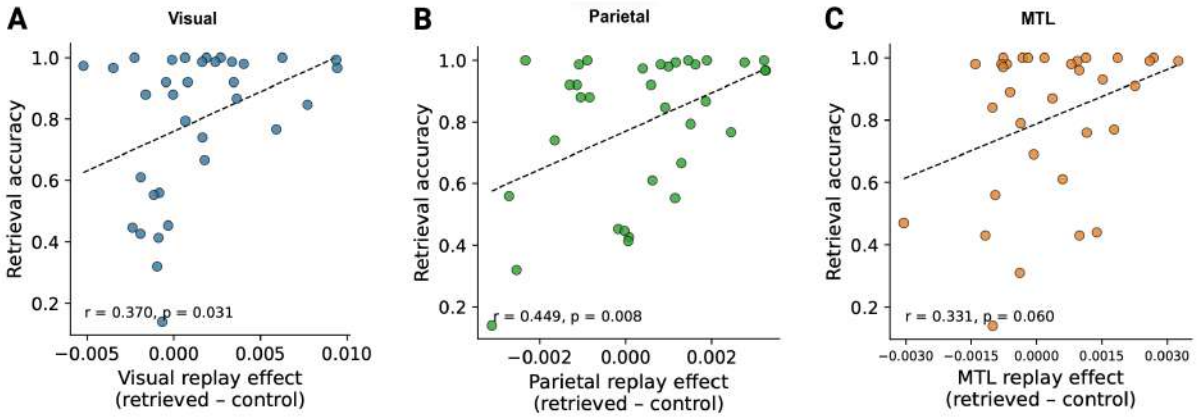

Figure 6.6: **Brain–behavior correlations for object retrieval.** **(A)** Subject-level correlation between visual cortex replay difference (absolute SODA, retrieved − control) and mean object retrieval accuracy ($r = 0.370$, $p = 0.031$). **(B)** Parietal replay difference vs. retrieval accuracy ($r = 0.449$, $p = 0.008$). **(C)** MTL replay difference vs. retrieval accuracy ($r = 0.331$, $p = 0.060$). Note that using the raw retrieved signal (not subtracting control) yields a significant association ($r = 0.372$, $p = 0.030$; see main text). Each dot represents one participant; dashed lines show linear fits. $N = 34$.

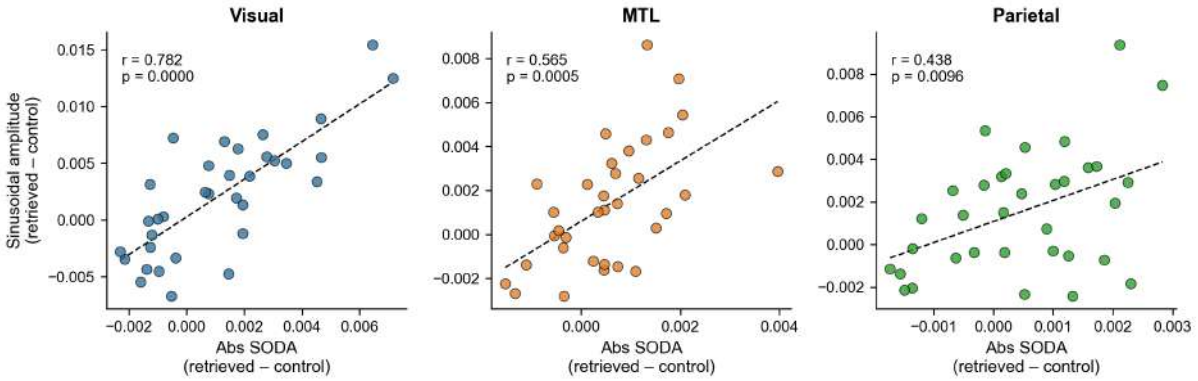

Figure 6.7: **Convergence of replay metrics.** Subject-level correlation between the absolute SODA replay effect (retrieved − control) and SinAmp replay effect (retrieved − control) for visual cortex (left), MTL (center), and parietal cortex (right). Strong positive correlations in all three ROIs confirm that the two metrics capture consistent individual differences in sequential reactivation strength. Each dot represents one participant. $N = 34$.

6.3 Discussion

6.3.1 Summary of findings

These results establish five key findings. First, the hierarchical sequence memory task produces robust behavioral performance with graded difficulty across hierarchical levels and serial positions, and retrieval accuracy is preserved across sleep while reaction times decrease. Second, logistic regression classifiers trained on independent localizer data (combining encoding and

retrieval blocks) successfully decode individual stimulus identities from fMRI patterns across ROIs, generalize to the main task, and detect stimulus-specific reinstatement during retrieval even in the absence of visual input. Third, sequential reactivation during retrieval is detectable in all three ROIs as indexed by the absolute SODA metric and confirmed by SinAmp. When split by session, the regional profile shifts: visual cortex carries the signal before sleep, while MTL and parietal cortex emerge after sleep. Fourth, parietal replay for objects increases across sleep, as shown by a region-specific retrieval $\times$ session $\times$ ROI interaction that strengthens in the trial-matched analysis. The parietal signal is concentrated in the first post-sleep retrieval block and decays by the second block. Fifth, brain–behavior correlations link replay to memory at the subject level: participants with stronger visual cortex and parietal replay show better object retrieval accuracy, and MTL replay predicts accuracy particularly post-sleep. However, replay does not predict accuracy at the trial level, likely reflecting limited power given the strong imbalance between correct ($\sim$85%) and incorrect ($\sim$15%) trials. Notably, the dissociation between scene and object replay effects maps onto the hierarchical structure of the task: scenes, which serve as higher-order contextual anchors and are retrieved near ceiling, show no session-dependent changes, whereas objects, which carry the bulk of behavioral variance, show the sleep-dependent parietal emergence. A fuller analysis of how representational structure at different hierarchical levels changes across consolidation will be part of the forthcoming publication.

6.3.2 Interpreting the replay metric: sequential reactivation versus general reinstatement

A critical question for the interpretation of the results concerns the extent to which the observed replay effects reflect genuinely sequential reactivation as opposed to general reinstatement of retrieved items. The shuffled-label validation analysis bears directly on this issue.

We repeated the replay detection analysis using shuffled sequence position labels, which destroy the temporal ordering while preserving item-level information. If the absolute SODA metric captured purely sequential structure, shuffled labels should abolish the effect entirely. Instead, many shuffled comparisons remained significant (11/18 tests at $p < 0.05$), indicating that the metric captures a composite signal: it reflects both genuine sequential ordering *and* general reinstatement strength: higher overall activation for retrieved versus control items, regardless of their temporal position. This is important for interpretation: the replay effects reported here could in principle be driven by elevated reactivation of individual retrieved items, which would produce a larger SODA slope even without genuinely sequential ordering.

Critically, the SinAmp metric suffers from the same limitation. Any slope in the SODA time course, whether generated by sequential or non-sequential reinstatement, will produce a periodic signal that the sinusoidal model can fit. Indeed, the SinAmp for the true sequence was not larger than for shuffled sequences. The current data therefore do not provide sufficient evi-

dence to conclude that the signal is fully sequential. Rather, the effects should be understood as reflecting preferential reinstatement of the retrieved sequence, with sequentiality remaining a plausible but not exclusively supported component. Importantly, the finding that this preferential reinstatement reorganizes from visual to parietal cortex across sleep constitutes evidence for sleep-dependent changes in the neural basis of retrieval, regardless of whether the underlying signal is fully sequential.

This finding motivates ongoing methodological work aimed at dissociating the sequential and non-sequential components of the replay signal. The correlation between the absolute SODA and SinAmp metrics (Fig. 6.7) confirms that both measures capture consistent individual differences, but future analyses will need to identify conditions or contrasts under which the two metrics diverge in order to isolate the unique contribution of temporal ordering.

6.3.3 Implications for systems consolidation

The sleep-dependent emergence of parietal sequential reactivation aligns with the core prediction of the complementary learning systems framework: that sleep-dependent replay drives the redistribution of memory representations from hippocampal to cortical networks [McClelland et al., 1995, Frankland and Bontempi, 2005]. The finding that parietal replay increased from pre-sleep to post-sleep, and that this effect was specific to object trials with their greater behavioral variance, suggests that consolidation selectively strengthens cortical representations for memories that benefit most from reorganization.

The emergence of parietal cortex as the locus of post-sleep reactivation is consistent with a growing literature implicating lateral parietal regions, particularly the angular gyrus and intraparietal sulcus, in episodic memory retrieval [Shimamura, 2011, Rugg and King, 2018]. The angular gyrus has been proposed to serve as a convergence zone that binds multimodal features into coherent episodic representations [Bonnici et al., 2016], and its engagement increases for richly detailed memories [Richter et al., 2016]. Within the systems consolidation framework, parietal cortex may thus be a natural target for the cortical integration of sequential structure: as replay during sleep strengthens cortical traces, parietal regions that support multi-feature binding during retrieval begin to carry the reactivation signal that was initially dependent on hippocampal and sensory cortices.

The MTL results are more nuanced. Rather than declining after sleep as a strict interpretation of systems consolidation might predict, MTL replay emerged post-sleep. This may reflect the continued role of the hippocampus in coordinating retrieval even after partial cortical takeover, consistent with models in which consolidation is a gradual process and hippocampal involvement persists during early stages of cortical integration [Sekeres et al., 2017, Barry and Maguire, 2019]. Indeed, recent evidence suggests that the hippocampus may never fully disengage from vivid episodic retrieval, but instead shifts its contribution from detailed reinstatement to a more schematic or indexing role [Moscovitch et al., 2016, Robin and Moscovitch, 2017].

6.3.4 Limitations and future directions

Several limitations should be noted. First, as discussed in the preceding subsection, the absolute SODA metric captures a composite of sequential ordering and general reinstatement strength. The shuffled-label validation demonstrated that the replay effects reported here cannot be attributed exclusively to temporal ordering; elevated reactivation of retrieved items, regardless of their sequential position, contributes to the signal. Developing analytic contrasts that isolate the unique contribution of sequentiality remains an important goal for future work.

Second, the pre-sleep session comprised one retrieval block while the post-sleep session comprised two. This power asymmetry is addressed by the trial-matched analysis, which restricts the post-sleep data to the first retrieval block and thereby equates trial counts across sessions. The fact that the parietal session effect strengthened in the trial-matched analysis indicates that the asymmetry does not inflate the post-sleep results; if anything, the additional post-sleep data dilute the effect, as the parietal signal decays across the second retrieval block. A closer examination of how sequential reactivation changes across the morning retrieval blocks will be part of the forthcoming publication.

Finally, the brain–behavior associations are between-subject: participants with stronger reactivation retrieve more items correctly. The absence of a trial-level accuracy effect most likely reflects a power limitation, given the strong imbalance between correct and incorrect trials.

6.3.5 Relation to earlier thesis chapters

The present findings complement the wakefulness replay results reported in Chapter 2. That chapter applied the same SODA framework to detect sequential replay during pre-decision pauses in a value-based learning task, demonstrating that on-task replay operates over inferred latent cause structure, with replay strength tracking non-local value updating on a trial-by-trial basis. The present chapter extends this line of work by applying SODA to memory retrieval across overnight consolidation, showing that replay-like signatures are detectable during active retrieval and that their regional profile shifts after sleep. Together, the two projects suggest that replay supports complementary functions across behavioral contexts: during on-task decision-making, it facilitates value propagation within learned latent structures; during post-sleep retrieval, it reflects the redistribution of memory representations toward cortical networks. An important difference between the two studies is the timescale of interest: Chapter 2 focused on rapid, within-trial replay dynamics linked to ongoing value computations, whereas the present chapter examines how the overall strength and locus of replay change across a night of sleep. Bridging these timescales and determining whether the same replay mechanisms that support online value propagation also contribute to overnight consolidation remains an important direction for future work.

6.4 Methods

6.4.1 Participants

Forty-six participants were recruited from the participant pool of the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig (mean age = 29.7 years, SD = 5.3; 21 females). All gave written informed consent. The study was approved by the ethics committee at the University of Leipzig (128/23-ek). Eleven participants were excluded at various stages due to technical difficulties, including incomplete MRI sessions, missing behavioral data, and absent localizer recordings. Because different analysis steps have different data requirements, each analysis includes the maximum number of participants with complete data for that specific pipeline: $N = 39$ for behavioral analyses (including participants with complete behavioral but partial MRI data), $N = 35$ for classifier cross-validation within localizer data, $N = 34$ for retrieval decoding analyses, and the encoding transfer analysis.

6.4.2 MRI acquisition

MRI data were acquired on a 3T Siemens Magnetom Prisma Fit system equipped with a 64-channel head/neck coil at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig.

Functional MRI (fMRI) scans were collected using T2*-weighted gradient-echo echo-planar imaging (GE-EPI) with multiband acceleration, sensitive to BOLD contrast. Functional images had an isotropic voxel size of $2 \times 2 \times 2$ mm, a repetition time (TR) of 1.5 s, an echo time (TE) of 28 ms, and a flip angle of 72°. Sixty-five slices were acquired in an interleaved order oriented approximately 20° toward the coronal plane (T > C 20°), with a field of view of 204 mm, a multiband acceleration factor of 5, 6/8 partial Fourier in the phase-encoding direction, and fat saturation. Phase encoding was in the anterior–posterior direction. The first five volumes of each block were discarded to allow for scanner equilibration.

High-resolution structural images were acquired for spatial normalization, coregistration, and cortical surface reconstruction. T1-weighted images were collected using a magnetization-prepared rapid gradient-echo (MPRAGE) sequence ($1 \times 1 \times 1$ mm isotropic; TR = 2300 ms; TE = 2.98 ms; TI = 900 ms; flip angle = 9°; 176 sagittal slices; FoV = 256 mm; GRAPPA factor 2). A T2-weighted image was acquired using a SPACE sequence ($0.9 \times 0.9 \times 0.9$ mm isotropic; TR = 3200 ms; TE = 408 ms; 192 sagittal slices; FoV = 230 mm; GRAPPA factor 2).

To correct for susceptibility-induced distortions in the functional data, field maps were acquired using spin-echo EPI sequences with opposing phase-encoding polarities (anterior–posterior and posterior–anterior; PEPOLAR method; $2 \times 2 \times 2$ mm isotropic; TR = 7000 ms; TE = 51 ms; 65 slices).

6.4.3 MRI preprocessing

Structural and functional MRI data were preprocessed using fMRIPrep 25.1.4 [Esteban et al., 2019]. Preprocessing included brain extraction, spatial normalization to native T1w space (for decoding analyses) and MNI152NLin2009cAsym standard space, head-motion correction (using `mcflirt`), susceptibility distortion correction (PEPOLAR), slice timing correction, and confound estimation. Confound time series included six rigid-body motion parameters and their temporal derivatives, global signal, framewise displacement, and six anatomical CompCor components (aCompCor), yielding 20 nuisance regressors. FreeSurfer 7.3.2 [Fischl, 2012] was used for cortical surface reconstruction and anatomical parcellation.

6.4.4 Region-of-interest definition

Regions of interest (ROIs) were defined from subject-specific FreeSurfer anatomical parcellations and transformed to native functional space. Three ROIs were analyzed:

- **Visual cortex** (VIS): 1005/2005, 1011/2011, 1021/2021, 1013/2013, 1007/2007, 1009/2009, 1015/2015
- **Medial temporal lobe** (MTL): 17/53, 1006/2006, 1016/2016
- **Parietal cortex**: 1008/2008

To mitigate signal dropout in regions susceptible to magnetic susceptibility artifacts, ROI masks were intersected with intensity-thresholded functional images. For each run, voxels with mean signal intensity below 30% of the global mean were excluded from the ROI mask.

6.4.5 Voxel reliability selection

To improve signal quality for multivariate analyses, a voxel reliability selection procedure was applied. Within each anatomical ROI, voxel selection was based on two criteria: activation strength and reliability. First, activation strength was quantified as the mean BOLD signal intensity across the entire time course, excluding voxels with poor signal quality or signal dropout. Second, voxel reliability was assessed using a split-half correlation approach [Tarhan and Konkle, 2020]. Beta estimates were derived from a run-wise first-level GLM implemented in Nilearn (SPM canonical HRF; cosine drift model; high-pass filter = 1/128 Hz; voxel-wise standardization; version 0.10.3). For each run, the design matrix included onsets and durations for all stimuli and, when present, a probe condition. Nuisance regressors comprised the global signal, framewise displacement, six motion parameters and their temporal derivatives (12 total), and six anatomical CompCor components (aCompCor00–05). Voxel reliability was then computed by correlating beta response profiles between odd and even localizer runs. For each voxel, condition-wise beta maps were averaged within odd and even sets to yield paired

response profiles across all regressors of interest. Pearson correlations between these profiles produced voxel-wise reliability maps. Voxels exceeding the 70th percentile reliability threshold were retained. To increase robustness, reliability masks from the two independent localizer sessions were intersected, retaining voxels that were reliable in at least one session. The resulting combined masks were used for all subsequent decoding analyses.

6.4.6 Multivariate pattern analysis: classifier training

Logistic regression classifiers were trained on the two independent localizer sessions to decode individual stimulus identities (10 scenes, 10 object categories) from fMRI activity patterns within each ROI. In each localizer session, all 60 task stimuli (50 objects and 10 scenes) were paired with unique, unrelated words. During encoding blocks, participants studied word–image pairs (3.5 s per pair, ∼2.5 s inter-trial interval). During retrieval blocks, words were presented alone (5 s per trial), requiring participants to internally recall the associated image and indicate retrieval success via button press with occasional probes (10%). Classifiers were trained on neural patterns evoked during both encoding and retrieval blocks, combining stimulus-driven responses (from image viewing) with retrieval-related patterns (from word-cued recall). The inclusion of retrieval trials was motivated by the goal of capturing neural patterns that more closely approximate the top-down memory retrieval process engaged during the main task, where participants must internally reconstruct sequence elements from positional cues.

Separate classifiers were trained for scenes and objects, and classifier performance was evaluated using stratified leave-one-run-out cross-validation within the localizer data.

6.4.7 Cross-session decoding transfer

Classifiers trained on localizer data were applied, without retraining, to neural activity recorded during the encoding session and the retrieval sessions (pre-sleep and post-sleep). For each trial, the classifier produced a probability time course for each stimulus category across a 10-TR window following event onset (∼15 s). This cross-session generalization approach tests whether category-specific representations established during localizer viewing are reactivated during memory encoding and retrieval.

6.4.8 Sequential slope ordering analysis

To assess whether memory retrieval reinstated the temporal order of encoded sequences, we applied the sequential slope ordering analysis (SODA) following Wittkuhn and Schuck [2021], Renz et al. [2026]. SODA exploits the temporal dynamics of overlapping hemodynamic responses: when a neural sequence is rapidly replayed, each element evokes a BOLD response offset in time. Due to the sluggish nature of the hemodynamic response, these overlapping responses create characteristic temporal patterns in classifier probability time courses. If stimuli

are replayed in order $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4 \rightarrow S_5$, the resulting classifier probabilities show ordered dynamics: initially, earlier items have higher probabilities than later items, creating a forward-ordered pattern; as the hemodynamic responses evolve, this pattern reverses. To quantify this ordering, for each retrieval trial classifier probabilities for the five sequence items were extracted at each TR, and at each time point a linear regression was fitted:

$$p_i(t) = \beta_0 + \beta_1 \cdot \text{position}_i + \varepsilon \tag{6.1}$$

where $p_i(t)$ is the decoded probability for the object at sequence position i at TR t, and β_1 is the SODA slope. A positive slope at a given TR indicates that later-sequence items currently have higher decoded probabilities, while a negative slope indicates the reverse. A sequential replay event thus produces a characteristic sign change in the SODA slope across time.

For each trial, SODA slopes were computed for both the *retrieved* sequence and a *control* sequence. For scene retrieval trials, the control was the scene sequence from the other context (e.g., night scenes when day scenes were probed). For object retrieval trials, the control was the object sequence drawn from the other category sequence within the same scene structure (e.g., F–J when A–E was probed, or vice versa). The difference between retrieved and control SODA values served as the primary measure of sequential reactivation. Two additional shuffle controls (random permutation and a selected permutation excluding identity, reverse, and adjacent-preserving orders) provided supplementary baselines.

6.4.9 Phase-invariant sinusoidal model fitting

Because the temporal onset of sequential retrieval varies across trials, averaging raw SODA slopes can lead to phase cancellation (see Results for a detailed motivation). To obtain a phase-invariant measure of sequential retrieval strength, a sinusoidal model was fitted to each trial's SODA time course:

$$y(t) = A \cdot \sin(2\pi f t + \varphi) + c \tag{6.2}$$

where A is the amplitude (retrieval strength), f is the frequency (in cycles per TR), φ is the phase, and c is a baseline offset. Model fitting was performed using nonlinear least-squares optimization (`scipy.optimize.curve_fit`) with frequency bounds $f \in [0.03, 0.45]$ cycles/TR and five random restarts to mitigate local minima. The fitted amplitude A (sinusoidal amplitude, SinAmp) served as the primary phase-invariant measure of sequential retrieval strength; the goodness of fit (R^2) was used as a secondary measure to validate the presence of periodic structure (see Results).

6.4.10 Statistical analysis

Unless otherwise noted, group-level comparisons of replay metrics were performed using paired t-tests (within-subject pre- vs. post-sleep) or one-sample t-tests (testing retrieved–control dif-

ferences against zero). Linear mixed-effects models (LMEs) were fitted using statsmodels with subject as random intercept and REML estimation, used to test session effects, ROI $\times$ session interactions, and behavioral data. Behavioral retrieval rates were modeled as subject-level percentages with session as fixed effect and subject as random intercept. Brain–behavior correlations were computed as Pearson correlations between subject-level neural metrics and behavioral accuracy. Where families of tests were performed across ROIs, sessions, or metrics, false discovery rate (FDR) correction was applied using the Benjamini–Hochberg procedure; both uncorrected and FDR-corrected p-values are reported in the text.

6.4.11 Software

All analyses were implemented in Python 3. Key packages included: nilearn [Abraham et al., 2014] for neuroimaging operations; scikit-learn [Pedregosa et al., 2011] for machine learning; imbalanced-learn for SMOTE resampling; SciPy [Virtanen et al., 2020] for optimization and statistical testing; statsmodels for mixed-effects modeling; and pandas/NumPy for data handling. fMRI data were preprocessed with fMRIPrep [Esteban et al., 2019] and FreeSurfer [Fischl, 2012].

Chapter 7

Neuroscience Methods for Investigating Brain Plasticity

Luianta Verra, Fabian Renz, Nicolas Schuck

Published in: *The Oxford Handbook of Cognitive Enhancement and Brain Plasticity*, Aron K. Barbey (ed.)
DOI: https://doi.org/10.1093/oxfordhb/9780197677131.013.24
Published: 23 September 2025

Abstract

The brain's ability to change in response to new environments, experiences, or damage—its plasticity—is a major foundation of cognition. Unraveling the mechanisms of plasticity in humans is therefore a central goal for neuroscience. This chapter provides an overview of common methods used to study structural and functional brain plasticity in humans. The main structural methods discussed are T1-weighted magnetic resonance imaging (MRI), which is widely used to track anatomical changes in the brain, and positron emission tomography (PET), which allows investigation of neurochemical processes, such as receptor density changes. The section on functional plasticity revolves around (T2*-weighted) functional MRI (fMRI) and attempts to provide an overview of analysis methods and MRI protocols that can be used to study functional changes in the brain, including resting state, univariate, and multivariate analyses. The chapter also discusses the dynamic nature of plasticity across different timescales, spanning from single days to several weeks. The text further provides a short summary of noninvasive brain stimulation methods that, combined with other structural or functional methods, have proven to be useful tools for inducing and measuring plasticity. The chapter also addresses limitations of these methods, for tracking fast or microscopic changes, and discusses the importance of a good study design that ensures that

measurements match the research question. This emphasizes the benefits of combining noninvasive brain imaging with longitudinal designs to reveal the nature of plasticity in humans.

Keywords: brain plasticity, cognition, neuroscience methods, magnetic resonance imaging, positron emission tomography, noninvasive brain stimulation

7.1 Introduction

Our brain has considerable capacity for change. It naturally evolves across development and aging and flexibly adapts to changing demands of the environment. The term *neural plasticity* has been used to describe these changes in the brain. Although definitions vary, most scientists agree that the term "neural plasticity" should be reserved for lasting changes that occur in response to experience or damage and have both structural and functional consequences [Lövdén et al., 2010b, Paillard, 1976]. For this chapter, we adopt a similar definition of plasticity based on Paillard (1976, translated in Will et al. 2008), according to which a change is plastic when it involves changes in brain structure that has functional consequences (i.e., for brain activity), conversely, functional changes that induce structural change. Following this definition, not all changes in the neural system are plastic. While brief activations, such as those caused by repeatedly perceiving an object, can trigger some forms of "short-term plasticity" [e.g., synaptic potentiation; Zucker and Regehr, 2002], these changes are only temporary, and we do not consider them in this chapter. Instead, we focus on the effects of long-term plasticity, such as long-term potentiation (LTP) or long-term depression [LTD; Malenka and Bear, 2004]. Structural adaptations in response to brain injuries, for instance, trigger lasting large-scale changes in brain structure and have consequences for cognitive function, and hence are an expression of plasticity. Similarly, learning a new language, childhood development, and aging all induce plasticity as the brain flexibly adapts to new demands over time. We note that in practice, it is often difficult to determine if an observed functional change results from structural changes in the brain or if the functional change precedes changes in brain structure. Analogously, observed changes in brain structures can be a cause or effect of functional changes. The methods we discuss in this chapter usually allow us to quantify functional or structural changes in the brain, but only the most carefully conducted experiments allow us to infer any direction or causality.

7.1.1 Timescale of Change

The ability to implement plasticity on a range of timescales is a highly desirable property for any system seeking to adapt to a complex world. A major factor determining the time scale of

change is the level of neural organization. Plasticity can, for instance, affect molecular signaling (e.g., receptor expression, changes in channel density), the cellular level (e.g., myelination), or networks of cells, and each of these levels places constraints on how dynamically it can be adapted. Yet while plastic changes are by definition considered "lasting," no general consensus on a minimum duration for a change to be considered plastic exists, with documented cases of plasticity ranging from hours to years [Miller and Constantinidis, 2024].

Broadly speaking, changes within single cells can occur in a very short timespan, but plasticity in cell networks or myelination reflects the cumulative effects of many single cell changes that may take days or even months to become apparent [Sampaio-Baptista et al., 2018]. Classic examples of long-term neural plasticity include the morphological changes observed in long-time musicians and the structural differences in the hippocampi of London taxi drivers [Gärtner et al., 2013, Maguire et al., 2000], which appear to develop over years of training. On the other end of the spectrum, it has been observed that brief neural high-frequency stimulation of a few seconds can induce changes lasting several hours or days within single cells [Bliss and Lomo, 1973]. This impressive potential for plasticity implies that the brain is constantly evolving in daily life, continuously forming new connections and reorganizing multiple interacting components. This makes it challenging to distinctly separate structural, plastic, and functional changes.

7.1.2 What We Cannot Measure (Directly) in Humans

Related to the questions about the levels at which plasticity occurs is how neuroscientists typically measure it, which poses particular challenges for human research. While this chapter largely focuses on human research, we briefly provide a summary of two core mechanisms of plasticity that have been studied mostly in animals, where neural processes can be probed invasively by directly observing individual nerve cells and where environmental, genetic, structural, and chemical factors influencing plasticity can be directly manipulated.

The first core mechanism is synaptic plasticity, which refers to changes in the strength or efficacy of synaptic transmission. The most common form of this process results from the temporal coincidence of presynaptic and postsynaptic neuronal spikes such that even brief millisecond-scale coactivations can lead to a longer-lasting potentiation or depression of synaptic connections, hence termed "spike-timing-dependent plasticity" [Caporale and Dan, 2008]. Such long-term potentiation and long-term depression effects have been investigated most extensively in the CA1 region in the hippocampus, where studies have shown that experimentally induced in vivo coactivations lead to changes in synapses and in turn induce long-term and spatial memory formation [see Citri and Malenka, 2008, for overview]. The second major mechanism that is difficult to study directly in humans is neurogenesis, a form of plasticity in which new neurons are generated from neural stem cells and integrated into existing neural networks. Neurogenesis occurs during development but also in the adult brain, although in this

case it is spatially restricted to a few regions, including the dentate gyrus region of the hippocampus [Boldrini et al., 2018, Eriksson et al., 1998]. To what extent neurogenesis persists across the life span is highly debated, but adult neurogenesis and its role in lifelong brain plasticity are exciting topics of investigation [Boldrini et al., 2018, Kempermann, 2022, Sorrells et al., 2018].

While synaptic plasticity and neurogenesis as observed in animal studies are the most foundational and best-understood aspects of plasticity, human research has focused on phenomena that reflect our natural circumstances and can be more easily studied. In the clinical domain, brain reorganization that occurs in response to trauma, for instance, is of great importance. Other commonly observed cases in humans in which brain plasticity plays a major role are more gradual and systemic. Prolonged exposure to stress, for instance, has been related to atrophy and loss of neurons in the prefrontal cortex and hippocampus [Duman et al., 2016]. Relatedly, aging is characterized by widespread plasticity, including synaptic alterations in areas such as the prefrontal cortex and the hippocampus [Morrison and Baxter, 2012].

Furthermore, practical and ethical constraints on human research require much more indirect approaches that rely predominantly on noninvasive recordings such as magnetic resonance imaging (MRI) and a range of methods based on it (see below). Because the MRI signal is spatially coarse, human research mostly reflects macroscopic changes in entire cell populations, unlike the synaptic plasticity and neurogenesis mechanisms in single cells at the center of animal research. Yet, as this chapter shows, recent advances in imaging techniques have allowed us to study neural reorganization in adult humans in vivo in unprecedented detail [Chen et al., 2010]. Hence, our overview of human work focuses on imaging studies of trauma, training, or environmental factors.

The remainder of this chapter is structured into five sections. First, we discuss design considerations that should be taken into account when attempting to measure plasticity, arguing that a good experimental design is essential to disentangle functional and structural components of plasticity. Next, we review methods that reflect structural brain changes, followed by an overview of functional methods. We then cover less common multimodal and neurochemical approaches. Finally, we provide some conclusions and future directions.

7.2 Design Considerations

Measuring plasticity and change inherently requires accounting for the temporal dynamics of a phenomenon of interest. Two research designs allow us to capture such dynamics to different degrees: longitudinal and cross-sectional designs. In longitudinal designs, the variable of interest is assessed repeatedly, such as before and after acquiring a new skill or at regular intervals during an intervention (e.g., training). Most importantly, the variable of interest is assessed in the same participants at each measurement point, allowing measurement of trajectories over time. Longitudinal studies are particularly suited for studying phenomena such as developmen-

tal changes but also for intervention studies such as drug trials or training studies. Alternatively, plasticity can be measured using cross-sectional designs, where individuals naturally varying in the variable of interest are assessed and compared. A cross-sectional design is often used to study changes that unfold over a long period of time and where a longitudinal design would be difficult to implement. The use of cross-sectional studies for studying change, however, has been criticized, given that their focus lies on measuring between-person differences rather than within-person change [Boker et al., 2009]. Inconsistencies between longitudinal and cross-sectional results in studies investigating cognitive aging, for instance, support this position [Lindenberger et al., 2011]. Here we briefly discuss the challenges and considerations of each design.

7.2.1 Longitudinal Study Designs

Longitudinal studies represent the gold standard in measuring plasticity as they allow us to learn about within-person change over time. However, designing and conducting longitudinal studies present multiple challenges. They are logistically demanding and require substantial resources. Additionally, various potentially confounding factors need to be accounted for, which include systematic dropout (e.g., fading motivation due to lack of improvement), changes in staff performing the study, or changes of setup (e.g., change of MRI scanner). Other confounds that potentially lead to biased conclusions are training and familiarity effects, meaning that improved performance over time might result from increased familiarity with the task rather than genuine improvement. When designing a longitudinal study, another consideration should be when and how often to measure in order to capture the change of interest [Hertzog and Nesselroade, 2003]. If a change unfolds over a long period of time, measuring in intervals of months or years might be sufficient. If change is rapid, measuring in yearly distances might fail to capture the trajectory of the plastic event. The Nyquist frequency (twice the frequency with which the phenomenon of interest changes) provides a useful lower bound for such design considerations.

7.2.2 Cross-Sectional Study Designs

While longitudinal designs are valuable because they provide insights into trajectories of change, cross-sectional designs allow us to measure differences between groups. In cross-sectional designs, groups hypothesized to differ in a specific process—such as musical experts versus novices, or older versus younger adults—are directly compared at a single time point. This approach is beneficial in cases where longitudinal designs are infeasible. For example, the neural changes elicited by learning to play an instrument can in principle be studied longitudinally. However, tracking participants over a timespan of years, while trying to account for confounding variables, is challenging. Cross-sectional approaches allow us to contrast expert musicians with a matched control group and to measure differences at a single time point. Such designs

have, for example, been used to show that musicians show differences in interhemispheric communication [Schlaug et al., 1995] or that young and older adults differ in the variability of the fMRI BOLD signal [Grady and Garrett, 2014].

However, cross-sectional designs come with their own set of challenges and limitations. Given that cross-sectional designs involve between-subject comparisons, measured change can be due to existing differences in the two groups rather than the variable assessed. One example of such differences is cohort effects, when observed differences between groups result from generational or cultural factors rather than from the variable of interest. In addition, age effects can further account for variability. Age effects include all differences that result from underlying psychological, social, or biological processes that highly correlate with age. A further critical consideration in cross-sectional designs is the causal directionality of an effect. For instance, do musicians develop different brain structures as a result of extensive training, or do individuals with certain neural characteristics become musicians because they possess the necessary predispositions? This question underscores the importance of careful experimental design and the use of complementary approaches to reveal the directionality.

Cross-sectional and longitudinal designs can be combined in what are known as *cross-sectional longitudinal designs*. This study design allows measuring groups of individuals of different ages over a period of time and conveniently combines cross-sectional with longitudinal elements. For example, Raz et al. [2005] combined a cross-sectional fMRI study with a longitudinal follow-up to investigate five-year changes in brain volume.

7.2.3 Choice of Control Group

One important consideration when attempting to measure plasticity, cross-sectionally or longitudinally, is the choice of control group. A *control group* is an additional group that does not receive the intervention of interest and thus contributes to causal understanding of an effect. Most importantly, control groups should be closely matched to the experimental group in all variables that are not of interest. The best way to equate possible a priori group differences is to offer so-called randomized control trials, where participants are randomly assigned to the intervention or control group [Hariton and Locascio, 2018]. Furthermore, choosing an appropriate control intervention for the control group is crucial: while applying no treatment at all is a common research strategy, it raises concerns that treatment effects could reflect unintended side effects of the treatment, such as frequent social interactions that accompany a study, or retest effects. To avoid such issues, researchers often use a placebo treatment or an alternative intervention unrelated to the variable of interest. For example, in a longitudinal study, Voss et al. [2010] assessed the effect of aerobic exercise in older adults on the brain's functional connectivity. Participants were randomly assigned to the experimental or control condition, with the control condition being an alternative exercise routine based on nonaerobic exercises. Another example is the use of sham stimulation in transcranial magnetic stimulation (TMS).

Given that stimulation with TMS is accompanied by a characteristic sound of the stimulation device and sensory side effects, control groups usually are administered a weak stimulation or receive stimulation of an alternative brain region that reproduces the sensory effects but that ideally does not interfere with the neural effects of the stimulation [Duecker and Sack, 2015]. Besides the choice of the control group, an additional consideration concerns experimenter effects, such as preferential treatment of one group over the other. In clinical trials, this is generally prevented by designing double-blind studies, where neither the participant nor the researcher knows which treatment a participant is receiving.

In summary, measuring change requires assessing within-person trajectories, and longitudinal designs with repeated measurement points are best suited for this purpose. Cross-sectional designs provide many advantages but have to be used keeping in mind that they provide snapshots of between-person differences without informing us about change over time. The choice of design primarily depends on the phenomenon of interest and is accompanied by additional considerations around the target population and feasibility. Additional considerations when choosing a suitable design concern the choice of the control group and potentially confounding effects such as age and cohort.

7.3 Methods for Measuring Plasticity

In the next sections, we provide an overview of common methods used to measure changes in brain structure, beginning with structural MRI. We then discuss two other structural MRI–based methods, diffusion tensor imaging (DTI) and voxel-based morphometry (VBM), which can be used to study changes in white matter and gray matter, respectively. Afterward, we focus on methods used to study functional change, which include functional MRI (fMRI) and positron emission tomography (PET). Table 7.1 provides an overview of the considered methods and their properties.

7.3.1 Measuring Structural Change

7.3.1.1 Structural MRI (T1-Weighted)

MRI is an imaging technique used widely to generate three-dimensional anatomical images of the brain or body in a number of ways (T1-weighted MRI is discussed here, and others are described below).

All MRI methods employ magnets that produce a strong magnetic field (B0) that forces the magnetic moment of hydrogen protons in the body to align with the field.[1] MRI is generally considered safe but given the strong magnetic field all metal objects need to be removed before scanning. Nonremovable metal bodies such as implants or pacemakers, therefore, are a contraindication for this technique. Further limitations of MRI scanning are restricted room and movement inside the scanner and noise levels during scanning.

Table 7.1: Overview of imaging methods discussed in this chapter. "What is measured" provides an overview of the signal measured while "Anatomical basis" describes the underlying physiological basis where a signal is believed to originate. The columns "Spatial scale" and "Data acquisition" describe the spatial resolution of the measure and the duration of signal acquisition, respectively. The values provided about time and space are approximate and should be understood as a general reference, as they are influenced by a number of factors, such as the specific hardware used for acquisition or ongoing technological developments.

Method	What is measured?	Anatomical basis	Spatial scale	Data acquisition	Invasiveness
Structural measures					
MRI-T1	Hydrogen molecule concentration	Tissue types (fat / water indicative of gray matter, white matter)	0.1–2 mm[(1)]	2–10 minutes	Noninvasive
DWI/DTI*	Diffusion of hydrogen molecules	White matter fiber tracts	1–5 mm[(2)]	4–15 minutes	Noninvasive
Functional measures					
PET	Radioactivity after radioactive tracer injection	Glucose metabolism / neurotransmitter density and activity, depending on tracer	3–10 mm[(3)]	30–45 minutes (plus waiting time for radiotracer absorption)	Intravenous/oral administration of radioactive tracer
fMRI*	Blood oxygenation	Metabolic demand triggered by neural activity	0.75–3 mm[(4)]	1–3.5 seconds per whole brain image	Noninvasive
TMS	TMS-induced changes in measure of interest (behavior / MRI)	Neural population activity (i.e., measured with EEG)	5–15 mm[(5)]	30–45 minutes (TMS) & 20–30 minutes (tDCS)	Magnetic stimulation

* MRI based methods; (1) e.g., Balchandani and Naidich 2015, Feinberg et al. 2023; (2) e.g., Holdsworth et al. 2019; (3) e.g., Catana 2019; (4) e.g., Uğurbil 2014, 2021; (5) e.g., Bolognini and Ro 2010, and Sliwinska et al. 2014.

So-called T1-weighted (T1W) MRI scans are widely used for assessments of brain structure, as they reliably indicate the relative amount of fat versus water of brain tissue in a given 3D location, known as a voxel.[2] Typical T1W MRI scans achieve a resolution on the order of 1.0 mm^3 (1 μl or 0.001 ml voxel volume), but recent developments of ultra-high-field-resolution brain imaging achieve even better spatial resolution [Balchandani and Naidich, 2015, Feinberg et al., 2023].

7.3.1.2 Voxel-Based Morphometry

A widespread analysis technique of T1W MRI data is voxel-based morphometry (VBM), in particular when researchers aim to examine changes in brain structure over time, or compare different groups of interest with high regional specificity. Unlike other analysis methods that focus on whole-brain volume or predefined regions of interest (ROIs), VBM emphasizes spatial resolution, providing a voxel-wise estimation of volume or gray matter density. This enables a detailed examination of structural differences across the entire brain and makes VBM a useful tool to study plasticity-induced structural brain changes. While typically used to assess differences or changes in gray matter density, it is also applicable to study changes in white matter, albeit with reduced sensitivity due to the homogeneity of large white matter regions [Ashburner and Friston, 2000, Kurth et al., 2015].

VBM has significantly advanced the understanding of brain structure, both in comparing different groups and in tracking plastic changes over time. For example, Sato et al. [2015] have used VBM to reveal gray matter volumetric differences among trained musicians when compared to hobbyists and nonmusicians, showing a positive correlation between the extent of musical training and gray matter volume. Another illustrative case is the structural adaptations associated with extensive navigation experience [Maguire et al., 2000]. Using VBM, studies on the brains of London taxi drivers have demonstrated an increased volume in the posterior hippocampi, a region associated with storing spatial representations. These findings suggest that the brain has a remarkable capacity for localized plastic changes in response to environmental demands.

7.3.1.3 Diffusion Weighted Imaging and Diffusion Tensor Imaging

Diffusion weighted imaging (DWI) is another MRI-based method that measures the diffusion of water molecules in the brain. Diffusion refers to the random Brownian motion of molecules. In open water, diffusion of water molecules is random and equally probable in all directions, but the structure of the human brain, in particular its fiber tracts, restricts such molecule diffusion. These tissue-dependent differences in diffusion create image contrast that allows MRI to visualize and quantify white matter tracts in the brain. DWI is applied in medical imaging, for example, in stroke patients [Baliyan et al., 2016].

Diffusion tensor imaging (DTI) refers to a specific type of modeling DWI datasets that an-

alyzes the three-dimensional shape of diffusion in space, known as the *diffusion tensor*. DTI provides information about the mean diffusivity in each voxel in addition to the directional preference of diffusion and diffusion rate. In white matter, the diffusion is less restricted along the axon and is directionally dependent (anisotropic). This property makes it particularly suitable for DTI modeling, allowing for the production of detailed images of white matter tracts.

DTI has been widely used as a clinical tool in the context of disorders that involve abnormalities in the brain's white matter and is increasingly used in research. It is particularly suitable to study change because it provides information about changes in the in vivo microstructure of the brain.

A study by Lövdén et al. [2010a], for example, used DTI in combination with a training intervention, involving memory and perceptual speed tasks, to investigate plasticity of white matter tracts in the anterior corpus callosum. In their study, they further compare younger and older adults and report training effects on white matter microstructures that were comparable in magnitude across age groups. DTI has also been used to study age-related decline in white matter tracts. Voineskos et al. [2012], for instance, report widespread age-related changes in diffusivity of cortico-cortical white matter fiber tracts.

DWI and DTI are valuable tools to study white matter changes in the brain. The main limitations of these approaches are shared with other MRI methods and include imaging artifacts as well as complex preprocessing and analysis methods [see Soares et al., 2013, for overview].

7.3.1.4 Positron Emission Tomography

Positron emission tomography (PET) is another approach that can be used to study structural plasticity in the brain. Although mainly used for clinical diagnosis of cancer, PET can be employed to study changes in blood flow and to track particular neurotransmitters in the brain [see Jones and Rabiner, 2012, for overview]. The main principle of PET is to first administer a radioactive tracer (often intravenously) that selectively binds to, for example, a neurotransmitter receptor or another protein, and in a subsequent PET scan detect the spatial distribution of the radioactive tracer in the brain. Common tracers in PET imaging are the glucose analogue 18F-FDG used to image glucose and oxygen-15 which provide an indirect measure of blood flow in the brain. FDG-PET studies, for instance, have revealed glucose metabolic reductions in various areas of the brain related to Alzheimer's disease, such as in areas of the cingulate cortex [Mosconi, 2005]. In addition to measuring metabolic changes, PET has been used to detect changes in transmitter activity over time, such as the age-related decrease in dopamine transporters in specific brain areas in healthy subjects [Volkow et al., 1996].

Another interesting application of PET is the imaging of synaptic density by using a ligand that binds to the synaptic vesicle protein SV2A. This approach has, for example, been used to show synaptic loss in the hippocampus in patients with Alzheimer's disease [Finnema et al., 2016] and also to investigate alterations in a broad range of neuropsychiatric disorders [Cai

et al., 2019].

PET imaging can be combined with MRI using specific PET/MRI acquisition systems. In such simultaneous acquisition of PET and MRI, PET can measure transmitter release, receptor occupancy, or glucose metabolism in addition to information about brain anatomy (MRI) and brain activity (fMRI) [Werner et al., 2015].

7.3.2 Measuring Functional Change

Functional magnetic resonance imaging (fMRI) builds on MRI technology and allows for measuring changes in brain activation through the blood oxygenation level dependent (BOLD) effect [Hillman, 2014, Ogawa et al., 1990]. The BOLD signal results from neural activity increasing oxygen consumption, which triggers a local increase in blood flow through mechanisms known as *neurovascular coupling*. This blood flow response causes an increase in (diamagnetic) oxyhemoglobin and a decrease in (paramagnetic) deoxyhemoglobin that leads to magnetic field distortions that fMRI can measure. While the BOLD signal is an indirect measure of brain activity [Poldrack, 2000], it offers great opportunities to study plasticity-induced functional changes noninvasively in the human brain.

7.3.2.1 Functional Connectivity and Resting State

Functional connectivity refers to similarities in brain activations and synchronization of brain regions and is assessed by measuring systematic coactivations between voxels, the spatial units of measurement in MRI. These coactivations build "functional" networks that are often closely aligned with anatomical structures, such as the default mode network [Guerra-Carrillo et al., 2014]. Generally we can distinguish between resting-state connectivity and task-based connectivity. Resting-state functional connectivity (rs-FC) measures spontaneous fluctuations in brain activity when a participant is not engaged in a specific task. To examine plasticity, rs-FC before and after an intervention are contrasted to capture experience-dependent plastic changes. A notable example is the study by Newbold et al. [2020], in which participants' dominant arms were immobilized in a cast for two weeks, inducing motor deprivation. Using daily resting-state fMRI scans, it was shown that cortical and cerebellar regions associated with the immobilized limb disconnected from the broader motor network within forty-eight hours. Interestingly, the internal connectivity within the unused subcircuit remained intact, and large spontaneous activity pulses were observed—suggesting a protective mechanism to preserve existing connectivity. Critically, these changes were reversible: after the cast's removal, the previously disconnected regions gradually reintegrated into the motor network, with functional connectivity returning to baseline levels within days. This study provides compelling evidence of the adult brain's plastic capacity and demonstrates how fMRI can effectively capture plastic changes in functional connectivity in the human brain.

In rs-FC studies, plastic changes are often tracked longitudinally, with pre- and postintervention comparisons over weeks or months [Guerra-Carrillo et al., 2014, Powers et al., 2012, Vahdat et al., 2011]. However, researchers increasingly explore ways to track plastic changes on shorter timescales, such as during learning or memory formation, within individual experimental sessions [Taren et al., 2017, Yuan et al., 2014].

In addition to resting-state measures, functional connectivity can be assessed while participants actively perform a task, using so-called task-based functional connectivity measures. In a study by Allegra et al. [2020], functional connectivity was assessed continuously while participants performed a perceptual decision task where they had to identify the dominant direction of motion of a random dot cloud. During the task, dot color became related to the response, meaning that participants could switch strategy to successfully solve the task. On-task connectivity measures revealed differences in connectivity when comparing periods of strategy optimization and strategy change, showing a clear example of short-term, on-task network plasticity.

A variety of analytical tools are used to contrast different experimental conditions or track changes in neural activation over time. For a more detailed overview, we point to the chapter on Network Neuroscience Methods for Investigating Brain Plasticity in this textbook.

7.3.2.2 Univariate Analysis

Univariate analysis is one of the most commonly employed methods for examining brain function and has been successfully applied to track changes in neural activation. For example, Moody et al. [2004] provided evidence of a compensatory shift in neural activity in Parkinson's disease patients. Specifically, as a result of diminished striatal function caused by their disease, patients exhibited increased recruitment of the medial temporal lobe (MTL) during an implicit learning task. This compensatory activation suggests an interaction between memory systems involving the MTL and striatum. Similarly, Poldrack et al. [2001] demonstrated a dynamic nature of memory system engagement during learning in healthy participants. Their findings revealed a shift from early-stage learning, which primarily engages the MTL, to later stages where MTL involvement decreases as the striatal system becomes increasingly recruited, reflecting a transition from declarative to procedural learning. Using univariate model–based fMRI, Schuck et al. [2015a] showed in a spatial navigation task that, compared to younger adults, older participants showed a shift toward more striatal activity that was related to processes linked to the MTL in younger adults. This suggests that a dynamic interplay of neural resources linked to memory is altered by age-related structural changes, in line with a wider line of research on such shifts in aging [Zahodne and Reuter-Lorenz, 2019]. In combination, these studies show how univariate techniques can reveal plasticity-induced changes in recruitment of neural resources—either compensating for impaired functionality or, in healthy participants, reflecting a dynamic shift in brain region engagement over the course of learning and memory formation.

A more recent approach for studying brain plasticity with fMRI is repetition suppression, also known as fMRI adaptation. This method capitalizes on the observation that neurons exhibit reduced responses to repeated presentations of the same stimulus, a phenomenon interpreted as a marker of increased neural efficiency [Grill-Spector and Malach, 2001]. The physiological basis for repetition suppression was first identified in primates, where neurons in the inferotemporal cortex showed decreased activity with repeated exposure to the same visual stimuli [Barron et al., 2016b]. This suppression effect, now documented across various brain regions and types of stimuli, reflects a general neural coding mechanism by which the brain optimizes its response to repeated information.

Repetition suppression has emerged as a pivotal method for examining how neural representations are fine-tuned through repeated exposure in human research. In an interesting study, Barron and colleagues used repetition suppression to track the formation of associated memories after a short learning session, and also showed that such markers of plasticity can disappear soon thereafter due to a process of inhibitory rebalancing, which presumably happens overnight [Barron et al., 2016a]. This study thereby demonstrates how functional measures can provide important insights into the complex nonlinear dynamics of plasticity, which previously also have been documented using structural measures [Wenger et al., 2017]. Repetition suppression has also revealed adaptations to the experienced structure of the environment that occur on short time scales. Garvert et al. [2023] observed a cross-stimulus enhancement effect—essentially an inverted form of suppression—emerging during a choice task, where this enhancement scaled with spatial distance between stimuli. This effect indicates that participants updated their cognitive maps in response to reward prediction errors encountered throughout the task.

7.3.2.3 Representational Similarity Analysis

A final set of analysis techniques applied to fMRI that has proven useful for studying plasticity comes from multivariate pattern analysis (MVPA) approaches.

Here, one prominent approach is representational similarity analysis (RSA), which compares activity patterns across conditions or time points by quantifying their similarity. A core strength of RSA is that by focusing on a neural similarity space, it abstracts from many details of the neural recording (e.g., timescales or how many units were recorded). This allows researchers to compare results across vastly different systems—for instance, between animal electrophysiology and human fMRI, or between human neural activation patterns and simulated computational models of brain function [Kriegeskorte et al., 2008].

An illustrative example of using RSA to track representational changes over time comes from Bellmund et al. [2019a], who investigated how the lateral entorhinal cortex (lEC) represents temporal structure. In this study, participants learned temporal and spatial relationships between objects encountered along a fixed route in a virtual city. By comparing multivoxel pattern similarity before and after the learning task, Bellmund et al. [2019a] demonstrated that

learning altered the neural representations of objects in lEC, such that they reflected the temporal structure of the task. This approach highlights how the entorhinal cortex evolves to map temporal structure in response to experience.

Another multivariate fMRI analysis technique is the application of classification algorithms that can detect the presence or strength of activity patterns associated with particular stimuli. Using this approach, Shibata et al. [2012] could, for instance, show that motion detection training leads to better classification of direction signals in the visual cortex. In another study, Atir-Sharon et al. [2015] have shown that pattern classifiers can be used to predict future performance in a semantic associative learning task, enabling the tracking of rapid plasticity processes that underlie such learning. In yet another study, Schuck et al. [2015b] used decoding to predict which participants would learn a visual-motor association before it was evident in behavior, indicating that MVPA can be used to identify plasticity processes.

In conclusion, fMRI provides a powerful method for studying brain plasticity, especially in longitudinal studies. Recent advances in techniques such as RSA, repetition suppression, and pattern classification have significantly expanded our ability to detect dynamic changes in neural activity patterns, offering deeper insights into core cognitive functions such as learning and memory.

7.3.3 Noninvasive Brain Stimulation: TMS and tDCS

The final key addition to the arsenal of methods for studying brain plasticity that we want to discuss is noninvasive brain stimulation techniques, such as transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS). These methods are not about measuring plasticity but rather can induce plasticity by interfering with and stimulating neural activity in the brain.

In TMS, a specific coil is used to produce a magnetic field that penetrates the scalp, inducing a flow of electric current that stimulates the underlying tissue [Bestmann, 2008]. Stimulation is generally limited to superficial areas of the cortex and can be of excitatory or inhibitory nature. The exact mechanism underlying TMS is, however, a matter of active debate [Siebner et al., 2022]. tDCS in contrast uses weak direct electrical currents that are continuously applied to the scalp between two sponge electrodes. tDCS is thought to change the resting membrane potential of neurons, causing lasting changes in cortical-evoked potentials.

TMS and tDCS can both be applied offline (before a specific task is administered) or online (during the task). Depending on the stimulated area, stimulation-induced changes in neural activity can have behavioral effects, such as changes in reaction times [Marshall et al., 2005] or motor responses [Saucedo Marquez et al., 2013]. In a study using TMS to investigate effects on language, for instance, Devlin and Watkins [2007] showed how TMS can induce so-called speech arrest when stimulating areas of the scalp corresponding to the left inferior frontal cortex. In this study, subjects were asked to count aloud and TMS stimulation resulted in subjects

being unable to proceed with counting [Devlin and Watkins, 2007, Pascual-Leone et al., 1991]. This shows how TMS can induce a "virtual" temporary lesion that is consequently expressed in behavioral changes and thus might help understand the plasticity processes that underlie real lesions.

To better understand the neural changes induced by TMS or tDCS, researchers often concurrently measure brain activations using imaging methods discussed above, such as EEG, PET, or fMRI. Koolschijn et al. [2019], for instance, used tDCS stimulation in combination with imaging and computational modeling to show how neural inhibition plays a role in protecting memories from interference. In their study, participants learned two overlapping but context-dependent memories and fMRI was used in combination with RSA analysis to measure the neural representations and behavioral interference of these two memories. Next, tDCS was applied to decrease the release of the neurotransmitter gamma-aminobutyric acid (GABA) in interneurons which has been associated with neural inhibition. The decrease of GABA was measured using noninvasive magnetic resonance spectroscopy (MRS). The authors found that tDCS-induced reduction of GABA was associated with more interference between two overlapping memories, indicating that GABA and neural inhibition play a role in gating learning and memory [Barron, 2021, Koolschijn et al., 2019]. Other examples of concurrent brain stimulation and functional imaging are a range of studies showing the effects of focal stimulation on neural network activity. Using TMS, for instance, Hartwigsen [2018] has shown that virtual lesions of specific brain areas lead to up- or downregulation of activity in other areas of a network—a phenomenon that has been interpreted as a flexible compensation mechanism [Hartwigsen, 2018].

Single stimulation sessions are thought to produce transient activation changes, lasting approximately one hour. In addition to short-term changes, more persistent changes have been reported for repeated stimulation [Monte-Silva et al., 2013]. Substantial variation in the responses to TMS or tDCS stimulation has been reported, highlighting the need for further research into factors undermining stimulation effectiveness [López-Alonso et al., 2014, Vergallito et al., 2022].

Recent developments further see the use of TMS and tDCS to treat chronic pain [Lefaucheur et al., 2008] or in language therapy to promote language recovery in poststroke aphasia [Hamilton et al., 2011, Hartwigsen and Volz, 2021].

7.4 Open Questions and Future Directions

The brain has an extraordinary capacity for reorganization in response to changing internal and external demands. This plasticity can be observed across a range of contexts, from spontaneous recovery following damage such as stroke or injury, to training-induced adaptations that occur over extended periods. Neural plasticity occurs across multiple levels of neural organization, ranging from microscopic molecular signaling changes to macroscopic adaptations in large-

scale neural networks. The timescales of these plastic changes vary significantly, from rapid changes within seconds or minutes on a microscopic level (such as LTP), to long-term structural adaptations and functional reorganization that unfold over hours, days, and years. While invasive animal studies can directly examine mechanisms such as synaptic plasticity and neurogenesis at a finely resolved scale in time and space, noninvasive imaging and analysis methods available in humans only access mostly macroscopic changes on slower timescales. Furthermore, they measure the biological processes indirectly—for instance, by tapping into magnetic changes that track alterations in blood flow. Despite their indirectness, modern-day techniques such as MRI and PET have provided unparalleled access to brain plasticity and the many ways in which it manifests in the brain and behavior. Over the past decades we have also witnessed a continued and significant improvement of our capacity to capture dynamic changes in the human brain. Here, we have sought to provide an overview of many recent developments in the study of neural plasticity in humans. We highlighted the diversity of analytical approaches that researchers have employed and continue to develop. Innovative approaches that have shed light on the multifaceted nature of plasticity include functional connectivity studies, VBM, RSA, or multivariate decoding techniques. We have also provided examples of studies that have integrated noninvasive brain stimulation methods, such as TMS and tDCS, with neuroimaging, showcasing the potential for studying plasticity by enabling temporary, reversible perturbations and causal inference about brain function. Pairing these techniques with refined analysis tools and improving neuroimaging should enable a better understanding of neural plasticity and support its clinical applications. Here we have covered some examples of studies that have used sophisticated designs to advance our understanding of plastic long term changes in humans. Newbold et al. [2020], for instance, used daily fMRI resting-state scans to show how restriction of arm movement can result in rapid reorganization of functional connectivity; Koolschijn et al. [2019] used a tDCS intervention in combination with magnetic resonance spectroscopy and fMRI representational similarity analysis to show how neural inhibition protects memories from interference.

We also highlighted that measuring within-person change requires careful consideration of the study design, which ultimately determines and limits what we can measure. We have highlighted in particular how the study of plasticity calls for repeated within-person measurements, and have provided examples where such designs have provided valuable contributions to our understanding of plasticity.

There are promising ongoing developments in the field that we have not covered in this chapter. Real-time neurofeedback (rtNFB) is one newly emerging technique, which allows researchers to provide live feedback to participants about their brain activity during the experimental session, with the goal to train the ability to one's own brain activity [Sitaram et al., 2017]. rtNFB has been applied to a variety of phenomena ranging from attention to memory and motor control among others [deBettencourt et al., 2015, Scharnowski et al., 2015]. Sulzer et al. [2013] provide an overview of the field, discussing progress and current limitations of

rtNFB. Another recent advancement is transcranial ultrasound stimulation (TUS), which has expanded the toolkit for noninvasive brain stimulation. TUS offers enhanced spatial precision, which makes it particularly promising for targeting specific neural circuits with greater accuracy. Emerging evidence suggests that TUS may promote neurogenesis, adding a potentially powerful approach for stimulating brain plasticity [Qin et al., 2024]. Another promising line of methodological advancement is the development of ultra-high-field MRI and combined PET-MRI. These technologies offer improved spatial resolution and greater detail on molecular and structural changes, thereby reducing the gap between invasive animal studies and noninvasive human research. Ultra-high-field MRI allows researchers to detect subtle structural changes in the brain that are associated with plasticity [Zsido et al., 2023], while PET-MRI enables simultaneous acquisition of structural and molecular data, providing more comprehensive insights into the biological underpinnings of plasticity. Noteworthy progress is also being made in neuroimaging analyses. Of particular interest are studies focusing on neural replay, the fast sequential reactivation of neural activity patterns that is thought to play an essential role in consolidation, and thereby plasticity [Wittkuhn et al., 2021]. Replay was originally identified in animal studies and long thought to be inaccessible with current techniques in humans. Thanks to advances in both magnetoencephalography [Liu et al., 2019] and fMRI [Schuck and Niv, 2019, Wittkuhn and Schuck, 2021] analysis techniques, replay is now increasingly studied in humans [Liu et al., 2022].

Notes

[1]Protons naturally spin on their axis, behaving like magnets themselves. When aligned with B0, they create a net magnetization called M0. For imaging, short radio-frequency pulses are applied to stimulate the protons, causing them to go out of phase and the magnetization to tip away from B0. MRI sensors then capture the energy released from protons once they realign with the magnetic field. During this process of relaxation, the magnetization at each time point can be described as the *recovered longitudinal magnetization along B0* (T1) and the *remaining transverse magnetization that decays exponentially to 0* (T2). Different tissues such as fat or water have different relaxation times, which in combination with a suitable pulse sequence, can be exploited to create image contrast.

[2]T1W MRI leverages differences in longitudinal magnetization times between fat and water.

Part IV

Synthesis

Chapter 8

General Discussion

This thesis has examined the role of neural replay in shaping and utilising the internal representations that support learning, memory, and decision-making. The first empirical project demonstrated that replay during wakefulness supports the propagation of value across inferred latent structure, linking trial-by-trial replay strength in visual cortex to non-local value updating. The second shifted the focus to sleep, showing that sequential reactivation during memory retrieval undergoes a region-specific reorganisation across overnight consolidation: visual cortex carried the sequential signal before sleep, and parietal cortex emerged as the dominant locus after sleep. Together, these projects converge on a central theme. Replay is not a passive echo of experience but an active computational process that builds, refines, and utilises the brain's internal models.

What follows takes a particular stance on how to interpret these findings. This discussion treats "replay" not as a single mechanism with one canonical function, but as an umbrella term for fast sequential activation, an implementational family whose most productive contemporary interpretation is as sampling from an internal generative model [Kurth-Nelson et al., 2023, Schwartenbeck et al., 2023, Dayan et al., 1995]. On this view, value updating, planning, memory consolidation, structure learning, and the construction of novel, never-experienced sequences are not separate mechanisms competing for explanatory priority, but different downstream consequences of the same sampling operation, shaped by task demands. A recent formal treatment by Sagiv et al. [2025] makes a version of this claim quantitatively explicit: the "value" and "map" hypotheses of replay can be recovered as limiting cases of a single prioritisation framework in which replay targets trajectories informative for current and possible future goals, weighted by the agent's beliefs about which goals are likely to matter. The umbrella framing does not dissolve the distinctions between replay's proposed functions, but it reorganises them. What differs across brain states and tasks is not the mechanism, but what the sampling is for.

The discussion is organised around two interconnected threads. The first concerns function: what does the sampling do in the empirical work of this thesis, and how do the two projects map onto the umbrella view just sketched? The second concerns detection methods: how do

we detect fast sequential activation in humans, and what are the consequences of our methodological choices for the conclusions we can draw? These threads are not independent. What we can conclude about the function of replay is constrained by how we measure it, and the theoretical questions we wish to answer should drive the development of better tools.

8.1 Overview of Contributions

This thesis comprises three contributions. The following overview characterises each in terms of its question, design, and main findings, before the two subsequent sections develop the theoretical and methodological implications in greater depth.

8.1.1 Contribution 1: Latent cause inference, value generalisation, and neural replay during wakefulness

The first empirical contribution combined fMRI with computational modelling to examine how humans discover hidden reward structure and how neural replay supports the propagation of value across that structure. The study enrolled 52 participants in a two-day fMRI experiment centred on a sequential bandit task in which three of four options shared a common latent reward source. Successful performance required inferring the latent correlation structure from experience, a problem formalised here as latent cause inference [Gershman et al., 2010, Gershman and Niv, 2015], and then generalising value changes from observed to unobserved but structurally linked options.

On the behavioural side, participants showed systematic improvement in reversal-trial generalisation across the session, rising from approximately chance-level performance early in training to above-chance performance in the final two blocks. A latent cause inference model fitted to individual choice sequences outperformed standard temporal-difference accounts, including clairvoyant variants with access to the true latent structure from trial one. The model captured not only group-level learning trajectories but also individual differences. Participants achieving higher overall accuracy were best described by models inferring fewer latent causes enabling value generalization, and the degree to which the model showed selective non-local updating during reversals predicted individual reversal performance.

The neural analyses used the sequential slope ordering analysis approach applied to 18-second inter-stimulus intervals distributed across the task. Several novel findings emerged. Backward replay in visual cortex was selective for unobserved sequences sharing a latent reward source, and was absent for unrelated sequences and for sequences corresponding to the stimuli just observed, ruling out stimulus-driven residual activation as a confound. Replay strength increased across the session in line with participants' accumulating structural knowledge. Trial-by-trial fluctuations in replay were best predicted by non-local value updating derived from the participant-specific LCI model, at a granularity that distinguished which unob-

served path was being updated on a given trial. A metric capturing structural inference rather than value propagation did not reach significance in the same analysis. Replay strength also predicted behavioural accuracy over subsequent trials, positioning replay as prospective rather than merely reflecting past experience.

On the second day, the latent structure was covertly altered mid-session by reassigning one option to the previously independent reward source. This manipulation revealed a striking directional reorganisation of replay: associations that had been valid on Day 1 continued to elicit backward replay, whereas the newly valid generalisation links showed reliable forward sequentiality after the change, and a direct pre-versus-post comparison confirmed that this directional shift exceeded the pre-change baseline. The LCI model continued to account for trial-by-trial replay fluctuations on Day 2, and a representational similarity analysis of medial temporal lobe activity showed that MTL pattern geometry aligned progressively with the shared reward structure during Day 1, remained elevated at the start of Day 2, and reorganised following the structural change.

Taken together, these results provide the first direct evidence linking latent cause inference, as a formal model of structure discovery, to neural replay measured with fMRI, showing that replay is tightly coupled with the ongoing inference process in a manner consistent with a non-local value-updating function.

8.1.2 Contribution 2: Sleep-dependent reorganisation of sequential reactivation during memory retrieval

The second empirical contribution examined whether the regional profile of sequential memory reactivation during retrieval changes across a night of sleep, as the systems consolidation framework would predict. Thirty-nine participants (34 providing complete neural data) completed a hierarchical sequence memory task in which objects were organised into ordered sequences embedded within scenes and contexts, creating a three-level memory structure. Retrieval was assessed by fMRI both before and after overnight sleep, permitting a within-participant comparison of how the neural basis of retrieval reorganises across consolidation.

Behaviourally, retrieval accuracy was stable across sessions, with a modest 2–3 percentage-point improvement and a larger and more reliable decrease in reaction times from evening to morning. This relative behavioural stability is theoretically important, any session-related changes in neural signatures cannot be attributed to large shifts in memory strength and instead implicate representational reorganisation.

The central finding concerns the regional profile of sequential reactivation measured by SODA. Collapsing across sessions, robust sequential reactivation was detectable in visual cortex, medial temporal lobe, and parietal cortex, with effects surviving false-discovery-rate correction for both the absolute SODA metric and a complementary sinusoidal amplitude measure developed to address phase cancellation across trials. When sessions were examined sepa-

rately, a dissociation emerged: before sleep, only visual cortex showed significant sequential reactivation. After sleep, parietal cortex emerged as a significant locus while visual cortex no longer reached significance. Parietal cortex showed a region-specific retrieval-by-session interaction that strengthened in a trial-matched analysis equating trial counts across sessions and was confirmed by both metrics after multiple-comparison correction. The parietal signal was concentrated in the first post-sleep retrieval block and decayed over the course of the second. MTL sequential reactivation also emerged post-sleep, though this effect did not survive correction in the trial-matched analysis, leaving open the question of whether it reflects genuine consolidation-dependent change or a power advantage of the larger post-sleep dataset.

Brain–behaviour correlations linked replay to memory at the individual-differences level. Participants with stronger visual cortex replay and stronger parietal cortex replay retrieved more objects correctly, with both associations surviving false-discovery-rate correction. MTL replay showed a trend toward the same association, which strengthened post-sleep. A shuffled-label validation analysis established that the absolute SODA metric captures a composite of sequential ordering and general reinstatement strength, motivating interpretive caution. The session-dependent regional shift constitutes evidence for changes in how and where memories are retrieved after sleep, but whether the underlying signal is driven by genuine temporal ordering or elevated non-sequential reinstatement of retrieved items remains an open question addressed in the limitations.

Methodologically, this study introduced the sinusoidal amplitude metric as a phase-invariant complement to the trial-averaged SODA slope, motivated by simulation work revealing that phase cancellation across trials with variable onset timing can severely attenuate or eliminate trial-averaged signals. These simulation results, reported in Thread 2 below, emerged directly from applying the SODA method to a retrieval context and constitute an independent analytical contribution of the thesis.

Together, the sleep project provides the first application of SODA to sequential memory retrieval across overnight consolidation.

8.1.3 Contribution 3: Neuroscience methods for investigating brain plasticity

The third contribution is a methodological book chapter, published in *The Oxford Handbook of Cognitive Enhancement and Brain Plasticity* [Verra et al., 2025], which provides a systematic overview of the noninvasive tools available for studying structural and functional brain plasticity in humans. The chapter covers T1-weighted MRI and positron emission tomography as the principal structural methods, discusses fMRI analysis approaches ranging from univariate activation contrasts to resting-state connectivity and multivariate representational analyses, and reviews noninvasive brain stimulation techniques as tools for both inducing and probing plasticity. It also addresses the timescales over which plastic changes are observable, the limits

of current methods for detecting fast or microscopic changes, and the design considerations necessary for valid inferences about experience-dependent neural change.

8.2 Thread 1: The Functional Architecture of Replay

8.2.1 Reframing replay: an umbrella for sampling from a generative model

A useful way to reconcile the diversity of proposed replay functions is to recognise that they may all reflect different read-outs of a common underlying generative model. Such a model is best understood as an internal, probabilistic account of how latent causes and structural regularities in the environment give rise to observable experience, and it can be run "forwards" to sample or simulate trajectories, states, or configurations that need not have been directly experienced [Dayan et al., 1995]. Under this view, replay is a mechanism for drawing samples from such a model [Kurth-Nelson et al., 2023]. Value updating, planning, consolidation, and structure learning can then be cast as different uses of the same sampling machinery, propagating reward signals through simulated trajectories, evaluating candidate futures, training cortical representations on model-generated experience, or assembling known elements into novel configurations to derive new knowledge [Behrens et al., 2018, Wittkuhn et al., 2021, Schwartenbeck et al., 2023].

This unifying stance is to some extent already present in classical computational accounts. In the DYNA architecture [Sutton, 1991], value updating and planning are linked. Simulated experience is fed into the same learning rule that updates values during direct experience, and the updated values guide subsequent choices. Whether we describe a given replay event as "updating values" or as "planning" depends on whether we focus on the learning rule applied to the replayed transition or on the behavioural consequence of the updated policy. Similarly, the complementary learning systems framework [McClelland et al., 1995] predicted that interleaved offline replay during sleep would both transfer information to cortical networks (consolidation) and extract statistical regularities across episodes (structure learning), as two effects of the same interleaved training regime.

More recent accounts push this unification further. The generative replay framework of Schwartenbeck et al. [2023] and the compositional computation view of Kurth-Nelson et al. [2023] both treat the hippocampus as a generative model capable of sampling novel sequences by recombining learned elements. On this view, what we describe as "consolidation" (the gradual emergence of abstract representations in the cortex), "planning" (the evaluation of hypothetical trajectories), "structure learning" (the extraction of compositional rules), and "gap-filling" (the inference of transitions that were never directly experienced but are implied by learned structure) may all be downstream effects of a generative sampling process.

Sagiv et al. [2025] have recently made a version of this unification formally explicit. They

extend Mattar and Daw [2018]'s prioritised replay framework to a setting in which the locations of rewards are unknown or dynamic, and show that the "value" hypothesis (replay to guide current choices) and the "map" hypothesis (replay to build and maintain an abstract representation of the environment) are special cases of a single prioritisation rule. The rule weighs trajectories by their expected usefulness for current and possible future goals, so the balance between goal-directed and map-building replay depends on the agent's beliefs about which goals are likely to become relevant. This reconciles the "paradoxical" observation that replay sometimes lags rather than leads behavioural learning, and can even be biased toward no-longer-preferred outcomes after reward revaluation [Carey et al., 2019, Gillespie et al., 2021], and it provides a quantitative handle on when replay should prioritise value propagation and when it should prioritise the maintenance of a more general, successor-like representation of the environment, which Sagiv et al. [2025] formalise as a Geodesic Representation.

8.2.2 Two read-outs of one generative model

Under the umbrella view, the two empirical projects of this thesis can be read as complementary sampling processes serving different functional purposes.

The first project (Chapter 5) instantiates sampling for value updating. The trial-by-trial alignment between backward replay and the non-local value updates prescribed by the participant-specific LCI model places this work alongside the MEG findings of Liu et al. [2021b], who showed that backward replay at reward receipt preferentially targeted non-local paths. What the present results add is that the underlying inferential machinery, here a latent cause model rather than a fixed task graph, is itself updated online, and that replay content reorganises directionally when the inferred structure changes. Sampling, in other words, is tied to the current state of the generative model, not to the agent's recent sensory history.

The second project (Chapter 6) instantiates sampling at retrieval through a generative model that has been reorganised by overnight consolidation. The post-sleep emergence of parietal sequential reactivation aligns with the systems-consolidation prediction of hippocampal-to-cortical redistribution [Born et al., 2006, McClelland et al., 1995, Brodt et al., 2023], and with the role of parietal cortex in multi-feature binding during episodic retrieval [Shimamura, 2011, Rugg and King, 2018]. The overall pattern is, however, only partially consistent with a strict reading of this framework. The MTL effect that emerged post-sleep but did not survive trial-matching admits two interpretations: a statistical artefact of the session-asymmetric power, or a genuine persistence of hippocampal engagement during early cortical integration, consistent with accounts in which the hippocampus does not fully disengage following consolidation [Sekeres et al., 2017, Moscovitch et al., 2016]. Distinguishing these possibilities requires analyses beyond what is reported here and will be taken up in the forthcoming publication.

Both empirical projects detected sequential replay in visual cortex, but this commonality should not be overinterpreted. Replay is likely generated centrally and expressed across

whichever cortical regions represent the replayed content [Kurth-Nelson et al., 2023, Wittkuhn et al., 2021]. Because both tasks used visual object stimuli and classifiers were trained on stimulus-evoked patterns in ventral visual areas, visual cortex is simply where decoding has the highest sensitivity to replay of these particular representations. What matters for interpretation is not the cortical site but what modulates the signal and which task demands are active.

The two projects also differ in the behavioural context in which sampling was measured. In the first, replay occurred without any retrieval demand and was prospective in character. In the second, replay was measured during effortful cued recall and was concurrent: stronger reactivation accompanied better retrieval within the same session. This dissociation maps onto a distinction between two modes of sampling. Prospective updating, which propagates value for future decisions, and reconstructive retrieval, which reinstates encoded sequences to support current memory access.

The two projects converge on what may be the most fundamental empirical observation of this thesis. Replay content is shaped by the agent's goals and how the internal model is leveraged to achieve them, rather than by recent sensory history. In the first project, the sensory control elicited no post-sensory replay; in the second, sequential reactivation was specific to retrieved sequences relative to matched controls.

8.3 Thread 2: Detecting Replay in Humans

8.3.1 The fundamental measurement problem

The study of neural replay in humans faces a measurement challenge that is absent in rodent electrophysiology. In rodents, simultaneous recordings from large populations of place cells provide direct access to individual replay events [Wilson and McNaughton, 1994, Foster, 2017, Davidson et al., 2009]. In humans, no currently available non-invasive method can achieve this level of temporal and spatial precision simultaneously. fMRI provides millimetre-scale spatial resolution across the whole brain but samples the BOLD signal on a timescale of seconds, temporally blurring events that unfold on the millisecond scale. MEG and EEG offer millisecond temporal resolution but with limited and uncertain spatial localisation, particularly for subcortical sources such as the hippocampus. Source localisation techniques applied to MEG and EEG data, as well as spectral analyses of SODA slope timecourses in fMRI, attempt to mitigate these respective limitations but cannot fully overcome them.

8.3.2 TDLM

Temporally delayed linear modelling (TDLM), introduced in Chapter 2, has produced some of the most compelling evidence for human replay to date. Liu et al. [2021b] used it to detect two physiologically distinct replay signatures at the moment of reward receipt: a fast forward

replay likely reflecting local experience encoding, and a slower backward replay that scaled with normative utility and preferentially represented non-local paths. Schwartenbeck et al. [2023] extended TDLM to detect generative replay during compositional reasoning, showing that replay sequences predicted participants' eventual solutions. A recent methodological investigation by Kern et al. [2026] begins to explore how TDLM performs outside these relatively favourable, task-locked regimes. They collected MEG during an 8-minute resting state immediately after participants had learned a graph of object associations to criterion, applied a pipeline closely following Liu et al. [2019, 2021b], and found no sequenceness at any time lag, in either direction, and no correlation with subsequent memory. Because an empirical null is inherently ambiguous, they complemented this with a hybrid simulation in which real neural patterns extracted from the localiser were inserted into a control resting state at varying densities, providing a more direct window onto the sensitivity of the pipeline. Two observations from that simulation highlight important open questions for the interpretation of TDLM results. First, sequenceness only surpassed permutation significance at relatively high replay densities, on the order of one event per second, substantially above the sharp-wave ripple rates typically reported for human intracranial resting state. Second, sequenceness values exhibited systematic, participant-specific baseline offsets in the control resting state that could be comparable to or larger than the sequenceness increase induced by inserted replay at plausible densities, raising the possibility that brain–behaviour correlations in resting-state analyses may partly reflect baseline variability rather than the replay signal of interest. A bootstrap power analysis suggested that relatively large samples may be needed to reliably detect such correlations. Importantly, these findings do not undermine the TDLM evidence obtained in task-locked designs with strong within-trial contrasts, where sequenceness is computed over short windows in which brief high-density replay bursts are plausible. They do, however, point to a set of challenges that the field will need to address as it moves toward applying TDLM in more unconstrained settings. Establishing plausible replay-rate priors, validating pipeline sensitivity against realistic hybrid simulations, and developing approaches that can better disentangle baseline variability from genuine replay signals are all promising directions that could strengthen the interpretability of resting-state TDLM analyses going forward.

8.3.3 SODA

The sequential slope ordering analysis (SODA) method, developed by Wittkuhn and Schuck [2021] for fMRI data, takes a complementary approach. SODA exploits the fact that the slow hemodynamic response acts as a temporal integrator. When items in a sequence are reactivated in rapid succession, their overlapping BOLD responses create a systematic ordering in classifier-decoded probabilities that can be read out at each fMRI time point. At each TR, sequence position is regressed onto decoded probabilities, the resulting slope indexes the degree and direction of sequential reactivation.

A key advantage is the available spatial resolution: fMRI provides whole-brain coverage at millimetre resolution, enabling the study of where replay occurs and how it is coordinated across brain regions. This is precisely the information needed to disambiguate the signal origin separating cortical from subcortical replay signals. Wittkuhn and Schuck [2021] demonstrated that SODA can detect sequential structure at speeds far faster than the fMRI sampling rate. Wittkuhn et al. [2025] subsequently linked SODA-detected backward replay in visual cortex to implicit SR learning, and Schuck and Niv [2019] showed that hippocampal fMRI signals during post-task rest contain sequential structure consistent with replay of nonspatial task states. However, SODA carries its own set of assumptions and failure modes, which we have started to explore through simulation work.

8.3.4 What simulations reveal about SODA

To characterise SODA's assumptions and failure modes, we ran a set of targeted simulations probing how the detection metrics respond under varying assumptions about replay timing, classifier behaviour, and statistical inference. The following four observations are directly consequential for how fMRI replay studies should be designed, analysed, and interpreted.

Phase cancellation degrades trial-averaged signals

If replay events occur at variable times within the analysis window across trials, as they almost certainly do during unconstrained rest or self-paced retrieval, averaging raw SODA slope timecourses across trials produces destructive phase cancellation. Our simulations show that even modest jitter, on the order of a single fMRI volume, severely degrades the amplitude of the trial-averaged signal (Figure 8.1). A closely related problem arises within a trial when multiple replay events co-occur: their slope signatures combine constructively or destructively depending on their relative timing (Figure 8.2), and when forward and backward events overlap within the same analysis window their opposing sinusoidal signatures can partially or fully cancel, rendering both invisible to the SODA metric. This is not a purely theoretical concern. Liu et al. [2021b] demonstrated that forward and backward replay co-occur at the moment of reward receipt, albeit at different speeds, and there is no obvious reason to assume that forward and backward events do not co-occur during rest or retrieval as well. Both failure modes are rooted in the same problem, trial averaging assumes that the phase of the signal is stable across the units being averaged, and when it is not, contributions cancel. A natural remedy is therefore to compute summary metrics at the single trial level, where the phase structure of each event is still intact, and only then aggregate across trials. This finding directly motivated the phase-insensitive analysis strategy adopted in the second empirical project of this thesis. Rather than averaging signed slope timecourses, we used two complementary measures: the absolute SODA slope, which prevents positive and negative phases from cancelling one another, and a sinusoidal amplitude (SinAmp) obtained by fitting a sinusoidal model to each trial's

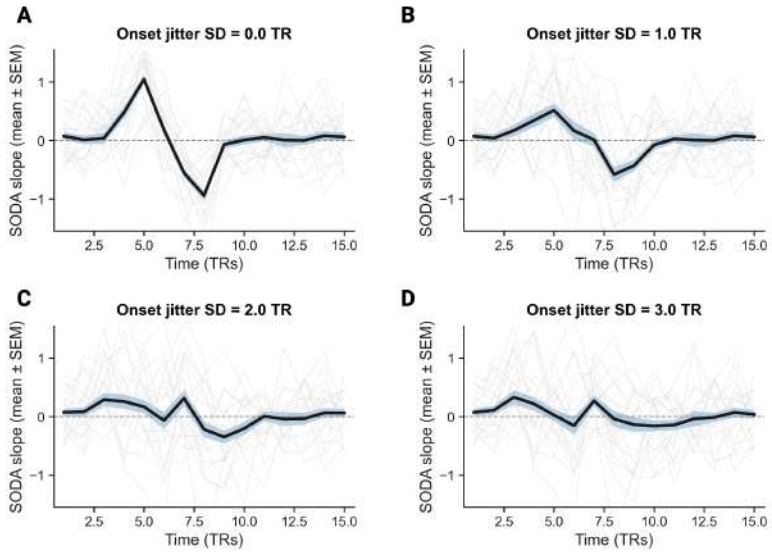

Figure 8.1: **Onset jitter degrades trial-averaged SODA signals.** Trial-averaged SODA slope timecourses (mean ± SEM, bold black line; individual trials in light grey) under increasing across-trial onset jitter. **(A)** With no jitter (SD = 0 TR), the trial-averaged slope produces a clean biphasic sinusoidal signature. **(B)** Modest jitter (SD = 1 TR) already substantially attenuates the amplitude. **(C, D)** Larger jitter (SD = 2–3 TRs) eliminates the trial-averaged signal almost entirely, even though individual trials still contain the underlying sequential structure.

SODA time course. The general implication is that any analysis relying on averaging decoded timecourses across trials with variable event timing or mixed directionality risks dramatically underestimating the true prevalence and strength of replay.

Heterogeneous classifier performance distorts the SODA waveform and biases directionality

SODA's slope metric assumes that all items in a sequence are decoded with approximately equal accuracy. In practice, classifiers rarely achieve uniform performance across stimulus classes, and our simulations show that even modest gradients in classifier sensitivity across sequence positions distort the expected shape of the SODA timecourse. For a genuine sequential reactivation event, the SODA slope approximates a sinusoid whose positive and negative half-cycles have roughly equal amplitude. When classifier sensitivity is uneven across positions, these two half-cycles take on different amplitudes, and the resulting asymmetric sinusoid is systematically biased toward the direction in which the accuracy gradient runs: if early-position items are decoded more accurately, the positive ("forward") half dominates, and if late-position items are decoded more accurately, the negative ("backward") half dominates (Figure 8.3). Heterogeneous classifiers thus do not generate spurious replay signals, but might distort directional interpretation. Without careful counterbalancing, any claim about forward versus backward replay drawn from SODA can be shaped by a residual accuracy gradient in

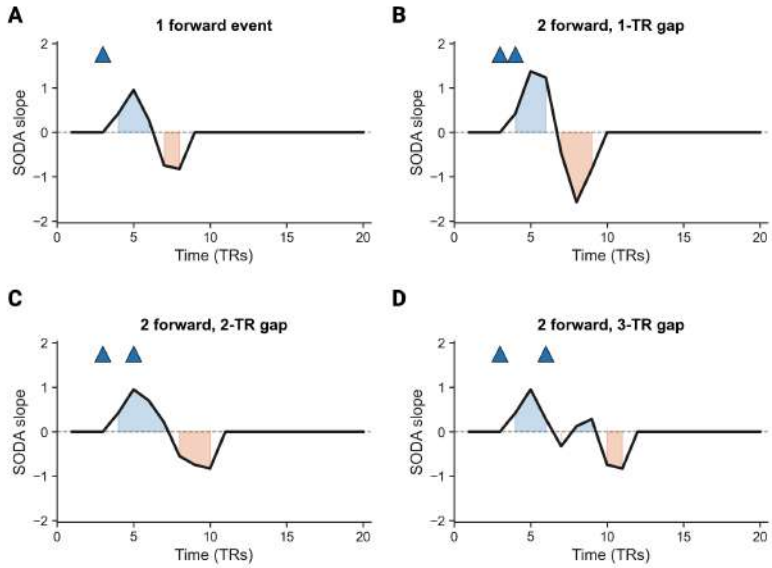

Figure 8.2: **Multiple forward replay events within a single trial can either reinforce or cancel one another depending on their relative timing.** SODA slope timecourses for **(A)** a single forward event and **(B–D)** two forward events separated by a 1-TR, 2-TR, and 3-TR gap respectively (event onsets marked by blue triangles; positive slope phases shaded blue, negative phases shaded orange). At a small inter-event gap (B), the positive half-cycles of the two events overlap and reinforce one another, producing a combined signature larger in amplitude than the single-event case. At a 2-TR gap (C), the constructive amplification largely cancels out and the combined signature is comparable in amplitude to a single event. At a 3-TR gap (D), the positive half-cycle of the second event lands on the negative half-cycle of the first, producing partial destructive cancellation: the combined signature is smaller in amplitude than a single forward event, even though two events are present. A combination of forward and backward events would complicate this picture further still, since opposing sinusoidal signatures from oppositely directed events can attenuate or eliminate one another entirely.

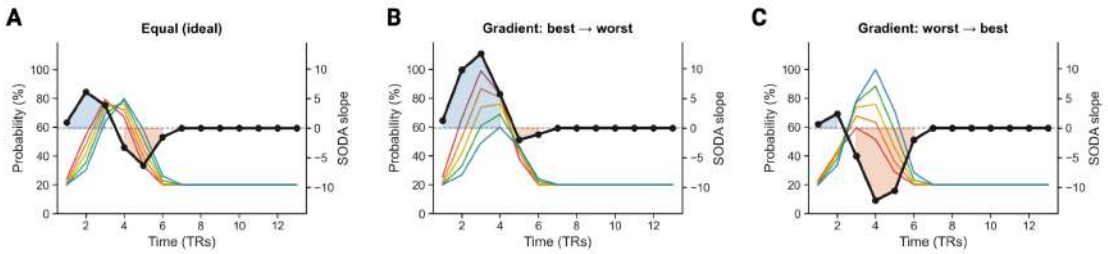

Figure 8.3: **Heterogeneous classifier performance distorts the SODA waveform and biases its directionality.** Per-class probability timecourses (coloured lines, left axis) and the resulting SODA slope (black line, right axis) for three classifier-amplitude configurations. **(A)** Equal amplitudes across all sequence positions: the SODA slope is a symmetric sinusoid. **(B)** Decreasing accuracy gradient (best at the start, worst at the end): the forward (positive) half-cycle dominates, biasing the metric toward apparent forward replay. **(C)** Increasing accuracy gradient (worst at the start, best at the end): the backward (negative) half-cycle dominates, producing the opposite bias. The shape distortion is purely a property of the classifier amplitude gradient, not of the underlying neural sequence.

the underlying classifiers. The most reliable solution is at the design level, counterbalancing sequence positions across stimuli so that any accuracy gradient averages out across conditions.

Z-scoring does not fix the heterogeneity bias

A natural methodological response to the heterogeneity bias described above is to rescale each classifier's probability timecourse before computing the slope, most commonly through per-class z-scoring within each trial. Our simulations show that this intuition bears its own caveats. At the single-trial level, z-scoring does remove the amplitude-induced shape asymmetry, and the slope timecourse recovers a symmetric forward / backward pair (Figure 8.4, panels B vs. E). However, when the weakest classifier in the set drops to or near the noise floor, z-scoring divides its pure noise by its own small within-trial SD, promoting it to full weight in the slope regression and injecting spurious peaks into the post-signal window that are comparable in magnitude to the true sequential reactivation (Figure 8.4, panels C vs. F). Whether z-scoring helps or harms therefore depends not on the amount of classifier heterogeneity but on whether the weakest classifiers still carry signal above their own noise floor. Furthermore, the absence of a probability peak need not reflect poor classifier accuracy; it may simply indicate that a particular state was not reactivated on that trial. Because the tested sequence is an experimenter decision, and SODA evaluates all states in that sequence jointly, z-scoring a flat, uninformative timecourse inflates its noise to the same scale as genuine reactivation peaks, thereby amplifying noise and introducing spurious replay signal.

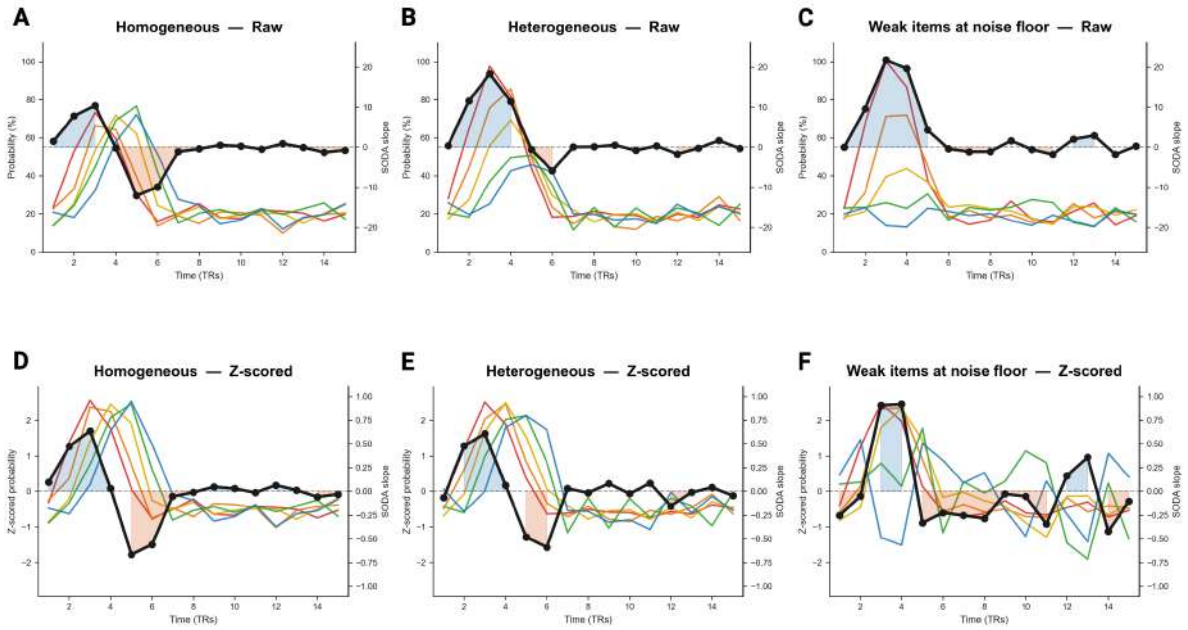

Figure 8.4: **Per-class z-scoring re-symmetrises the SODA waveform but contaminates the post-signal window when any classifier approaches the noise floor.** Single-trial classifier probability timecourses (coloured lines) and the resulting SODA slope (black line) under raw **(A–C)** and z-scored **(D–F)** pipelines, for three classifier-amplitude configurations: homogeneous **(A, D)**, heterogeneous with all items above the noise floor **(B, E)**, and weak items at the noise floor **(C, F)**. **A vs D:** both pipelines produce symmetric slope waveforms. **B vs E:** raw SODA shows an asymmetric signature biased by the amplitude gradient, while z-scoring re-symmetrises the sinusoid at the cost of discarding amplitude information. **C vs F:** z-scoring divides the pure noise of the weakest classifiers by their own small within-trial SD, promoting it to full weight in the slope regression and producing spurious peaks in the post-signal window comparable in magnitude to the true sequential reactivation.

Permutation nulls are not exchangeable

A natural way to ask whether an observed SODA slope reflects genuine sequential reactivation is to compare it against a null distribution constructed by permuting sequence labels. However, this strategy is not straightforward. The SODA slope is sensitive to the overall linear trend in decoded probabilities and therefore responds even to partial orderings, such as a sequence that is almost correct but with one pair of adjacent items swapped. Because of this, random permutations of sequence labels do not produce a uniform null distribution. The permuted set is dominated by near-correct orderings with substantial slope values, and the true ordering sits on a fat tail rather than on a sharply separated peak. Significance tests built on such permutation nulls are therefore overly conservative, and the absence of a significant effect should not be read as the absence of sequential reactivation. The practical recommendation is to prefer well-matched control conditions, for example pre-task rest or control sequences built from the same items in a different learned order, over naive permutation testing, and to treat marginally significant permutation-based results with caution. However, comparing against a control sequence composed of items drawn from a different, non-reactivated set introduces a different problem: a significant slope may be driven by the reactivation of individual items rather than by genuinely sequential replay, and therefore might not necessarily represent the full sequence, as discussed in Chapter 6. Evidence of sequentiality in the target sequence over a control condition is therefore necessary but not sufficient for establishing genuine replay, as it can also arise from the reactivation of individual items alone.

8.4 Limitations and Open Questions

Several questions lie beyond what the current generation of non-invasive methods and task designs can resolve. Rather than framing these as defects of the present work, the points that follow identify boundary conditions on what fMRI-based replay research, and this thesis in particular, can currently say and what might be promising future directions.

8.4.1 Dissociating structure learning from value updating

A central ambiguity runs through the first empirical project. The latent cause inference model was used to derive both structure-learning and value-updating signals, and both are driven by the same experimental events: reward reversals that simultaneously carry information about the value of individual sequences and the latent structure. The finding that trial-wise replay strength was better explained by non-local value updates than by a structure-learning metric is suggestive, but the close temporal coupling between structural inference and value propagation means that the two signals are difficult to dissociate statistically. It remains possible that replay contributes to structure learning in ways that are not captured by the metrics used here, or that

both processes draw on replay but with different temporal profiles or neural substrates that the current design cannot resolve. Future work could address this by experimentally decoupling structural change from value change, for example through paradigms in which the latent cause structure is altered without accompanying reward reversals.

8.4.2 Scope and interpretive constraints of the sleep project

The second empirical project reports neural analyses based on a substantial sample, and the central pre-sleep versus post-sleep comparison of sequential reactivation has been completed with robust results. The parietal emergence effect survives multiple-comparison correction in the trial-matched analysis and is confirmed by two independent metrics. Nevertheless, several constraints should be noted. The pre-sleep session comprised one retrieval block while the post-sleep session comprised two, creating a power asymmetry. The trial-matched analysis, which restricts the post-sleep data to the first retrieval block, addresses this concern and in fact strengthens the parietal effect, but the asymmetry means that the MTL post-sleep result (which emerged only in the full-data analysis) should be interpreted cautiously. Furthermore, as discussed in detail in the sleep chapter, the shuffled-label validation demonstrated that the absolute SODA metric captures a composite signal reflecting both sequential ordering and general reinstatement strength. The regional shift from visual to parietal cortex across sleep therefore constitutes evidence for sleep-dependent changes in the neural basis of retrieval, but whether the underlying signal is fully sequential or partly driven by elevated non-sequential reinstatement of retrieved items remains an open question.

8.4.3 The MTL puzzle

A finding that recurs across both empirical projects is the complex and partially unexpected behaviour of the medial temporal lobe. In the first project, MTL showed broad, relatively non-selective replay across all sequence conditions, in contrast to the sequence-specific backward replay detected in visual cortex. In the second project, MTL replay emerged post-sleep rather than declining as a strict interpretation of systems consolidation might predict, and its significance was limited to the full-data analysis (not surviving in the trial-matched comparison), raising the possibility that the MTL post-sleep effect reflects the greater statistical power of two post-sleep retrieval blocks rather than a genuine consolidation-related increase.

Two interpretations, not mutually exclusive, may account for the MTL pattern across the two studies. First, classifier sensitivity in MTL was substantially lower than in visual cortex in both empirical projects, which may constrain the granularity with which sequence-specific replay can be resolved, obscuring selectivity that is genuinely present in the underlying neural activity. Second, the MTL may play a qualitatively different role, for example providing a structural scaffold or indexing signal that drives more selective replay in cortical regions [Kurth-Nelson et al., 2016], rather than expressing sequence-specific content itself. This in-

terpretation is consistent with the first project's finding that MTL replay correlated with visual cortex replay across all conditions, and with the second project's observation that the region-specific session effect was carried by parietal cortex rather than MTL. If the signals detected in cortex are read-outs of hippocampally driven signals, then the spatial localisation provided by fMRI tells us where replay is expressed but not necessarily where it is generated.

8.5 Outlook

The two threads of this discussion converge on what this thesis specifically establishes and on the open questions that remain. Across the two empirical projects, the common functional picture is that replay is structured jointly by the agent's inferred model of the environment and by the functional goal it serves, rather than by recent sensory exposure. In the first project, backward replay in visual cortex tracked inferred latent cause structure, aligned with trial-by-trial non-local value updating rather than structural inference per se, and reorganised its content and direction when the latent structure changed. In the second, the regional profile of sequential reactivation during active memory retrieval shifted from visual to parietal cortex across overnight sleep, with parietal replay predicting retrieval accuracy at the individual-differences level. These findings add to a growing literature establishing that sequential reactivation in the human brain reflects task structure and value, is modulated by learning stage and sleep, and predicts subsequent memory and decision quality [Liu et al., 2021b, Wittkuhn et al., 2025, Schuck and Niv, 2019]. They also underscore that interpreting the signals we record remains constrained by methodological choices whose consequences are only beginning to be systematically characterised. Four directions seem particularly consequential for the next generation of replay research.

First, moving from detection to manipulation. The most compelling evidence for the causal role of replay in consolidation comes from targeted memory reactivation (TMR) studies [Rasch et al., 2007, Siefert et al., 2024, Geva-Sagiv et al., 2023], but TMR manipulates memory reactivation at the level of whole items or categories rather than at the level of specific sequential patterns. Developing methods to selectively enhance or disrupt sequential replay, for example through closed-loop neurostimulation triggered by decoded replay events, would provide the causal tests needed to move beyond the correlational evidence obtained in this thesis [Kurth-Nelson et al., 2023, Girardeau et al., 2009].

Second, bridging the gap between replay and representational change. This thesis has examined the two in parallel rather than connecting them directly. The first empirical project failed to find a link between representational change captured by representational similarity analysis and replay. The second empirical project has so far focused on replay detection without examining representational change, which will be included in the upcoming publication. A natural next step is to ask whether the content of replay during sleep predicts the specific representational changes observed after sleep.

Third, developing better computational models. Current computational accounts of replay tend to focus on one function at a time (e.g. value updating, consolidation, or structure learning). The generative replay framework of Schwartenbeck et al. [2023] and the SR-based accounts [Russek et al., 2017, Wittkuhn et al., 2025] represent steps toward unification, and recent models of autonomous hippocampal–cortical interactions during sleep [Singh et al., 2022] and generative consolidation via variational autoencoders [Spens and Burgess, 2024] have begun to formalise how replay transforms representational content during consolidation. A fully integrated model that accounts for replay across brain states, learning stages, and task demands remains elusive.

Fourth, advancing how replay is detected robustly in human neuroimaging. Thread 2 has shown that both of the dominant detection methods rest on assumptions that are only beginning to receive systematic scrutiny [e.g., Wittkuhn and Schuck, 2021]. For TDLM, Kern et al. [2026] made several of these assumptions explicit and showed that detection at the group level requires high replay densities, and that sequenceness-behaviour correlations can be shadowed by systematic baseline offsets rather than driven by the replay one intends to detect. For SODA, our simulation work mapped the failure modes arising from onset jitter, classifier heterogeneity, and the non-exchangeability of permutation nulls, and showed that constructing appropriate control conditions for statistical testing is considerably more delicate than a naive label-permutation test implies.

In sum, the functional contributions of this thesis are twofold: the first empirical project established that replay during wakefulness tracks inferred latent cause structure and propagates value across it on a trial-by-trial basis. The second showed that the regional profile of sequential reactivation during retrieval shifts from visual to parietal cortex across overnight consolidation, with the parietal emergence consistent with the redistribution of memory representations toward associative cortex, though the MTL pattern does not fit straightforwardly within this framework, as the post-sleep MTL effect only survived in the full-data analysis and may reflect the power asymmetry between sessions rather than a genuine consolidation-related change. The methodological contributions presented simulation work that characterised failure modes of the SODA metric and motivated the phase-invariant analysis strategies employed in the sleep project, while placing both SODA and TDLM within a critical framework for human replay detection more broadly. Together, these contributions advance the case that replay is best understood as sampling from an internal generative model, and that what differs across brain states is not the sampling operation itself but what it is for.

Bibliography

D. Abel, D. Arumugam, L. Lehnert, and M. Littman. State abstractions for lifelong reinforcement learning. In *International Conference on Machine Learning*, pages 10–19. PMLR, 2018.

A. Abraham, F. Pedregosa, M. Eickenberg, P. Gervais, A. Mueller, J. Kossaifi, A. Gramfort, B. Thirion, and G. Varoquaux. Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8, 2014. ISSN 1662-5196. doi: 10.3389/fninf.2014.00014. URL https://www.frontiersin.org/articles/10.3389/fninf.2014.00014/full.

L. Acerbi and W. J. Ma. Practical bayesian optimization for model fitting with bayesian adaptive direct search. *Advances in neural information processing systems*, 30, 2017.

D. J. Aldous. Exchangeability and related topics. In *École d'Été de Probabilités de Saint-Flour XIII—1983*, pages 1–198. Springer, 2006.

M. Allegra, S. Seyed-Allaei, N. W. Schuck, D. Amati, A. Laio, and C. Reverberi. Brain network dynamics during spontaneous strategy shifts and incremental task optimization. *NeuroImage*, 217:116854, Aug. 2020. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2020.116854. URL https://www.sciencedirect.com/science/article/pii/S1053811920303402.

J. R. Anderson. *Learning and memory: An integrated approach*. John Wiley & Sons Inc, 2000.

J. L. Andersson, S. Skare, and J. Ashburner. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, 20(2):870–888, 2003. ISSN 1053-8119. doi: 10.1016/S1053-8119(03)00336-7. URL https://www.sciencedirect.com/science/article/pii/S1053811903003367.

D. Aronov, R. Nevers, and D. W. Tank. Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, 543(7647):719–722, 2017.

J. Ashburner and K. J. Friston. Voxel-Based Morphometry—The Methods. *NeuroImage*, 11 (6):805–821, June 2000. ISSN 10538119. doi: 10.1006/nimg.2000.0582. URL https://linkinghub.elsevier.com/retrieve/pii/S1053811900905822.

T. Atir-Sharon, A. Gilboa, H. Hazan, E. Koilis, and L. M. Manevitz. Decoding the Formation of New Semantics: MVPA Investigation of Rapid Neocortical Plasticity during Associative Encoding through Fast Mapping. *Neural Plasticity*, 2015 (1):804385, 2015. ISSN 1687-5443. doi: 10.1155/2015/804385. URL https://onlinelibrary.wiley.com/doi/abs/10.1155/2015/804385. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2015/804385.

B. Avants, C. Epstein, M. Grossman, and J. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41, 2008. ISSN 1361-8415. doi: 10.1016/j.media.2007.06.004. URL http://www.sciencedirect.com/science/article/pii/S1361841507000606.

J. J. Bakermans, J. Warren, J. C. Whittington, and T. E. Behrens. Constructing future behavior in the hippocampal formation through composition and replay. *Nature neuroscience*, pages 1–12, 2025.

P. Balchandani and T. P. Naidich. Ultra-High-Field MR Neuroimaging. *American Journal of Neuroradiology*, 36(7):1204–1215, July 2015. ISSN 0195-6108, 1936-959X. doi: 10.3174/ajnr.A4180. URL http://www.ajnr.org/cgi/doi/10.3174/ajnr.A4180.

V. Baliyan, C. J. Das, R. Sharma, and A. K. Gupta. Diffusion weighted imaging: Technique and applications. *World Journal of Radiology*, 8(9):785, 2016. ISSN 1949-8470. doi: 10.4329/wjr.v8.i9.785. URL http://www.wjgnet.com/1949-8470/full/v8/i9/785.htm.

H. Barron, T. Vogels, U. Emir, T. Makin, J. O'Shea, S. Clare, S. Jbabdi, R. Dolan, and T. Behrens. Unmasking Latent Inhibitory Connections in Human Cortex to Reveal Dormant Cortical Memories. *Neuron*, 90(1):191–203, Apr. 2016a. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.02.031. URL https://linkinghub.elsevier.com/retrieve/pii/S0896627316001689. Publisher: Elsevier BV.

H. C. Barron. Neural inhibition for continual learning and memory. *Current Opinion in Neurobiology*, 67:85–94, Apr. 2021. ISSN 09594388. doi: 10.1016/j.conb.2020.09.007. URL https://linkinghub.elsevier.com/retrieve/pii/S0959438820301343.

H. C. Barron, M. M. Garvert, and T. E. J. Behrens. Repetition suppression: a means to index neural representations using BOLD? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1705):20150355, Oct. 2016b. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2015.0355. URL https://royalsocietypublishing.org/doi/10.1098/rstb.2015.0355.

D. N. Barry and E. A. Maguire. Remote memory and the hippocampus: A constructive critique. *Trends in cognitive sciences*, 23(2):128–142, 2019.

D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton. Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, 108(18):7641–7646, May 2011. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1018985108. URL https://pnas.org/doi/full/10.1073/pnas.1018985108.

T. E. Behrens, T. H. Muller, J. C. Whittington, S. Mark, A. B. Baram, K. L. Stachenfeld, and Z. Kurth-Nelson. What is a cognitive map? organizing knowledge for flexible behavior. *Neuron*, 100(2):490–509, 2018.

Y. Behzadi, K. Restom, J. Liau, and T. T. Liu. A component based noise correction method (CompCor) for BOLD and perfusion based fmri. *NeuroImage*, 37(1):90–101, 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2007.04.042. URL http://www.sciencedirect.com/science/article/pii/S1053811907003837.

J. L. Bellmund, P. Gärdenfors, E. I. Moser, and C. F. Doeller. Navigating cognition: Spatial codes for human thinking. *Science*, 362(6415):eaat6766, 2018.

J. L. Bellmund, L. Deuker, and C. F. Doeller. Mapping sequence structure in the human lateral entorhinal cortex. *eLife*, 8:e45333, Aug. 2019a. ISSN 2050-084X. doi: 10.7554/eLife.45333. URL https://elifesciences.org/articles/45333.

J. L. S. Bellmund, W. De Cothi, T. A. Ruiter, M. Nau, C. Barry, and C. F. Doeller. Deforming the metric of cognitive maps distorts memory. *Nature Human Behaviour*, 4(2):177–188, Nov. 2019b. ISSN 2397-3374. doi: 10.1038/s41562-019-0767-3. URL https://www.nature.com/articles/s41562-019-0767-3.

I. M. Berwian, P. Hitchock, S. Pisupati, G. Schoen, and Y. Niv. Using computational models of learning to advance cognitive behavioral therapy. *Communications Psychology*, 3(1):1–11, 2025.

S. Bestmann. The physiological basis of transcranial magnetic stimulation. *Trends in Cognitive Sciences*, 12(3):81–83, Mar. 2008. ISSN 1364-6613. doi: 10.1016/j.tics.2007.12.002.

C. M. Bishop and N. M. Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

T. V. Bliss and T. Lomo. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology*, 232(2):331–356, July 1973. ISSN 0022-3751. doi: 10.1113/jphysiol.1973.sp010273.

S. M. Boker, P. C. M. Molenaar, and J. R. Nesselroade. Issues in intraindividual variability: Individual differences in equilibria and dynamics over multiple time scales. *Psychology and Aging*, 24(4):858–862, Dec. 2009. ISSN 1939-1498, 0882-7974. doi: 10.1037/a0017912. URL https://doi.apa.org/doi/10.1037/a0017912.

M. Boldrini, C. A. Fulmore, A. N. Tartt, L. R. Simeon, I. Pavlova, V. Poposka, G. B. Rosoklija, A. Stankov, V. Arango, A. J. Dwork, R. Hen, and J. J. Mann. Human Hippocampal Neurogenesis Persists throughout Aging. *Cell Stem Cell*, 22(4):589–599.e5, Apr. 2018. ISSN 19345909. doi: 10.1016/j.stem.2018.03.015. URL https://linkinghub.elsevier.com/retrieve/pii/S1934590918301218.

N. Bolognini and T. Ro. Transcranial Magnetic Stimulation: Disrupting Neural Activity to Alter and Assess Brain Function. *Journal of Neuroscience*, 30(29):9647–9650, July 2010. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1990-10.2010. URL https://www.jneurosci.org/content/30/29/9647. Publisher: Society for Neuroscience Section: Toolbox.

H. M. Bonnici, F. R. Richter, Y. Yazar, and J. S. Simons. Multimodal feature integration in the angular gyrus during episodic and semantic retrieval. *Journal of Neuroscience*, 36(20):5462–5471, 2016.

J. Born, B. Rasch, and S. Gais. Sleep to remember. *The Neuroscientist*, 12(5):410–424, 2006.

M. Botvinick and M. Toussaint. Planning as inference. *Trends in cognitive sciences*, 16(10):485–488, 2012.

M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, and D. Hassabis. Reinforcement learning, fast and slow. *Trends in cognitive sciences*, 23(5):408–422, 2019.

S. Brodt, S. Gais, J. Beck, M. Erb, K. Scheffler, and M. Schönauer. Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science*, 362(6418):1045–1048, 2018.

S. Brodt, M. Inostroza, N. Niethard, and J. Born. Sleep—a brain-state serving systems memory consolidation. *Neuron*, 111(7):1050–1075, 2023.

G. Buzsáki. Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, 25(10):1073–1188, 2015.

Z. Cai, S. Li, D. Matuskey, N. Nabulsi, and Y. Huang. PET imaging of synaptic density: A new tool for investigation of neuropsychiatric diseases. *Neuroscience Letters*, 691: 44–50, Jan. 2019. ISSN 03043940. doi: 10.1016/j.neulet.2018.07.038. URL https://linkinghub.elsevier.com/retrieve/pii/S0304394018305172.

S. A. Cairney, N. El Marj, B. P. Staresina, et al. Memory consolidation is linked to spindle-mediated information processing during sleep. *Current Biology*, 28(6):948–954, 2018.

N. Caporale and Y. Dan. Spike timing-dependent plasticity: a Hebbian learning rule. *Annual Review of Neuroscience*, 31:25–46, 2008. ISSN 0147-006X. doi: 10.1146/annurev.neuro.31.060407.125639.

A. A. Carey, Y. Tanaka, and M. A. van Der Meer. Reward revaluation biases hippocampal replay content away from the preferred outcome. *Nature neuroscience*, 22(9):1450–1459, 2019.

C. Catana. Development of Dedicated Brain PET Imaging Devices: Recent Advances and Future Perspectives. *Journal of Nuclear Medicine*, 60(8):1044–1052, Aug. 2019. ISSN 0161-5505, 2159-662X. doi: 10.2967/jnumed.118.217901. URL https://jnm.snmjournals.org/content/60/8/1044. Publisher: Society of Nuclear Medicine Section: The State of the Art.

H. Chen, J. Epstein, and E. Stern. Neural plasticity after acquired brain injury: evidence from functional neuroimaging. *PM & R: the journal of injury, function, and rehabilitation*, 2(12 Suppl 2):S306–312, Dec. 2010. ISSN 1934-1482. doi: 10.1016/j.pmrj.2010.10.006.

Z. S. Chen and M. A. Wilson. How our understanding of memory replay evolves. *Journal of Neurophysiology*, 129(3):552–580, 2023.

R. Ciric, W. H. Thompson, R. Lorenz, M. Goncalves, E. MacNicol, C. J. Markiewicz, Y. O. Halchenko, S. S. Ghosh, K. J. Gorgolewski, R. A. Poldrack, and O. Esteban. TemplateFlow: FAIR-sharing of multi-scale, multi-species brain models. *Nature Methods*, 19:1568–1571, 2022. doi: 10.1038/s41592-022-01681-2.

A. Citri and R. C. Malenka. Synaptic Plasticity: Multiple Forms, Functions, and Mechanisms. *Neuropsychopharmacology*, 33(1):18–41, Jan. 2008. ISSN 0893-133X, 1740-634X. doi: 10.1038/sj.npp.1301559. URL https://www.nature.com/articles/1301559.

A. G. E. Collins and M. J. Frank. Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, 120(1):190–229, Jan. 2013. ISSN 1939-1471, 0033-295X. doi: 10.1037/a0030852. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0030852.

A. O. Constantinescu, J. X. O'Reilly, and T. E. Behrens. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292):1464–1468, 2016.

A. C. Courville, N. D. Daw, and D. S. Touretzky. Bayesian theories of conditioning in a changing world. *Trends in cognitive sciences*, 10(7):294–300, 2006.

R. W. Cox and J. S. Hyde. Software tools for analysis and visualization of fmri data. *NMR in Biomedicine*, 10(4-5):171–178, 1997. doi: 10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

A. M. Dale, B. Fischl, and M. I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, 1999. ISSN 1053-8119. doi: 10.1006/nimg.1998.0395. URL http://www.sciencedirect.com/science/article/pii/S1053811998903950.

T. J. Davidson, F. Kloosterman, and M. A. Wilson. Hippocampal replay of extended experience. *Neuron*, 63(4):497–507, 2009.

N. D. Daw, S. J. Gershman, B. Seymour, P. Dayan, and R. J. Dolan. Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6):1204–1215, 2011.

P. Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural computation*, 5(4):613–624, 1993.

P. Dayan and Y. Niv. Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology*, 18(2):185–196, 2008.

P. Dayan, G. E. Hinton, R. M. Neal, and R. S. Zemel. The helmholtz machine. *Neural computation*, 7(5):889–904, 1995.

M. T. deBettencourt, J. D. Cohen, R. F. Lee, K. A. Norman, and N. B. Turk-Browne. Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience*, 18(3):470–475, Mar. 2015. ISSN 1546-1726. doi: 10.1038/nn.3940. URL https://www.nature.com/articles/nn.3940. Publisher: Nature Publishing Group.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.

L. Deuker, J. L. Bellmund, T. Navarro Schröder, and C. F. Doeller. An event map of memory space in the hippocampus. *elife*, 5:e16534, 2016.

J. T. Devlin and K. E. Watkins. Stimulating language: insights from TMS. *Brain*, 130(3):610–622, Mar. 2007. ISSN 0006-8950, 1460-2156. doi: 10.1093/brain/

awl331. URL https://academic.oup.com/brain/article-lookup/doi/10.1093/brain/awl331.

A. Dickinson and B. Balleine. Motivational control of goal-directed action. *Animal learning & behavior*, 22(1):1–18, 1994.

S. Diekelmann and J. Born. The memory function of sleep. *Nature reviews neuroscience*, 11 (2):114–126, 2010.

Y. Ding, S. L. S. Dunn, J. J. Sakon, Z. M. Aghajan, C. Duan, Y. Zhang, J. I. Berger, A. E. Rhone, K. V. Nourski, H. Kawasaki, M. A. Howard, V. P. Roychowdhury, and I. Fried. Reading specific memories from human neurons before and after sleep. *bioRxiv*, 2025. doi: 10.1101/2025.07.01.662486. Preprint.

A. Disouky, M. A. Sanborn, K. Sabitha, M. M. Mostafa, I. A. Ayala, D. A. Bennett, Y. Lu, Y. Zhou, C. D. Keene, S. Weintraub, et al. Human hippocampal neurogenesis in adulthood, ageing and alzheimer's disease. *Nature*, pages 1–10, 2026.

C. F. Doeller, C. Barry, and N. Burgess. Evidence for grid cells in a human memory network. *Nature*, 463(7281):657–661, 2010.

B. B. Doll, D. A. Simon, and N. D. Daw. The ubiquity of model-based reinforcement learning. *Current opinion in neurobiology*, 22(6):1075–1081, 2012.

F. Duecker and A. T. Sack. Rethinking the role of sham TMS. *Frontiers in Psychology*, 6, Feb. 2015. ISSN 1664-1078. doi: 10.3389/fpsyg.2015.00210. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2015.00210/full. Publisher: Frontiers.

R. S. Duman, G. K. Aghajanian, G. Sanacora, and J. H. Krystal. Synaptic plasticity and depression: new insights from stress and rapid-acting antidepressants. *Nature Medicine*, 22(3):238–249, Mar. 2016. ISSN 1078-8956, 1546-170X. doi: 10.1038/nm.4050. URL https://www.nature.com/articles/nm.4050.

D. Dupret, J. O'neill, B. Pleydell-Bouverie, and J. Csicsvari. The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature neuroscience*, 13(8): 995–1002, 2010.

S. J. Durrant, C. Taylor, S. Cairney, and P. A. Lewis. Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, 49(5):1322–1331, 2011.

J. M. Ellenbogen, J. C. Hulbert, R. Stickgold, D. F. Dinges, and S. L. Thompson-Schill. Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. *Current Biology*, 16(13):1290–1294, 2006.

J. M. Ellenbogen, P. T. Hu, J. D. Payne, D. Titone, and M. P. Walker. Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences*, 104(18):7723–7728, 2007.

J. M. Ellenbogen, J. C. Hulbert, Y. Jiang, and R. Stickgold. The sleeping brain's influence on verbal memory: boosting resistance to interference. *PLoS One*, 4(1):e4117, 2009.

P. S. Eriksson, E. Perfilieva, T. Björk-Eriksson, A.-M. Alborn, C. Nordborg, D. A. Peterson, and F. H. Gage. Neurogenesis in the adult human hippocampus. *Nature Medicine*, 4(11): 1313–1317, Nov. 1998. ISSN 1078-8956, 1546-170X. doi: 10.1038/3305. URL https://www.nature.com/articles/nm1198_1313.

O. Esteban, C. Markiewicz, R. W. Blair, C. Moodie, A. I. Isik, A. Erramuzpe Aliaga, J. Kent, M. Goncalves, E. DuPre, M. Snyder, H. Oya, S. Ghosh, J. Wright, J. Durnez, R. Poldrack, and K. J. Gorgolewski. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16:111–116, 2019. doi: 10.1038/s41592-018-0235-4.

A. Evans, A. Janke, D. Collins, and S. Baillet. Brain templates and atlases. *NeuroImage*, 62 (2):911–922, 2012. doi: 10.1016/j.neuroimage.2012.01.024.

D. A. Feinberg, A. J. S. Beckett, A. T. Vu, J. Stockmann, L. Huber, S. Ma, S. Ahn, K. Setsompop, X. Cao, S. Park, C. Liu, L. L. Wald, J. R. Polimeni, A. Mareyam, B. Gruber, R. Stirnberg, C. Liao, E. Yacoub, M. Davids, P. Bell, E. Rummert, M. Koehler, A. Potthast, I. Gonzalez-Insua, S. Stocker, S. Gunamony, and P. Dietz. Next-generation MRI scanner designed for ultra-high-resolution human brain imaging at 7 Tesla. *Nature Methods*, 20(12): 2048–2057, Dec. 2023. ISSN 1548-7091, 1548-7105. doi: 10.1038/s41592-023-02068-7. URL https://www.nature.com/articles/s41592-023-02068-7.

S. J. Finnema, N. B. Nabulsi, T. Eid, K. Detyniecki, S.-f. Lin, M.-K. Chen, R. Dhaher, D. Matuskey, E. Baum, D. Holden, D. D. Spencer, J. Mercier, J. Hannestad, Y. Huang, and R. E. Carson. Imaging synaptic density in the living human brain. *Science Translational Medicine*, 8(348):348ra96–348ra96, July 2016. doi: 10.1126/scitranslmed.aaf6667. URL https://www.science.org/doi/10.1126/scitranslmed.aaf6667. Publisher: American Association for the Advancement of Science.

B. Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.

V. Fonov, A. Evans, R. McKinstry, C. Almli, and D. Collins. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, Supplement 1:S102, 2009. doi: 10.1016/S1053-8119(09)70884-5.

D. J. Foster. Replay comes of age. *Annual review of neuroscience*, 40(1):581–602, 2017.

P. W. Frankland and B. Bontempi. The organization of recent and remote memories. *Nature Reviews Neuroscience*, 6(2):119–130, 2005.

P. W. Frankland, B. Bontempi, L. E. Talton, L. Kaczmarek, and A. J. Silva. The involvement of the anterior cingulate cortex in remote contextual fear memory. *Science*, 304(5672):881–883, 2004.

R. M. French. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences*, 3(4):128–135, 1999.

S. Gais, G. Albouy, M. Boly, T. T. Dang-Vu, A. Darsaud, M. Desseilles, G. Rauchs, M. Schabus, V. Sterpenich, G. Vandewalle, et al. Sleep transforms the cerebral trace of declarative memories. *Proceedings of the National Academy of Sciences*, 104(47):18778–18783, 2007.

C. R. Gallistel and J. Gibbon. Time, rate, and conditioning. *Psychological review*, 107(2):289, 2000.

M. M. Garvert, R. J. Dolan, and T. E. Behrens. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *elife*, 6:e17086, 2017.

M. M. Garvert, T. Saanum, E. Schulz, N. W. Schuck, and C. F. Doeller. Hippocampal spatio-predictive cognitive maps adaptively guide reward generalization. *Nature Neuroscience*, 26 (4):615–626, Apr. 2023. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-023-01283-x. URL https://www.nature.com/articles/s41593-023-01283-x.

S. J. Gershman. The successor representation: its computational logic and neural substrates. *Journal of Neuroscience*, 38(33):7193–7200, 2018.

S. J. Gershman and N. D. Daw. Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual review of psychology*, 68:101–128, 2017.

S. J. Gershman and Y. Niv. Learning latent structure: carving nature at its joints. *Current opinion in neurobiology*, 20(2):251–256, 2010.

S. J. Gershman and Y. Niv. Novelty and Inductive Generalization in Human Reinforcement Learning. *Topics in Cognitive Science*, 7(3):391–415, July 2015. ISSN 17568757. doi: 10.1111/tops.12138. URL https://onlinelibrary.wiley.com/doi/10.1111/tops.12138.

S. J. Gershman, D. M. Blei, and Y. Niv. Context, learning, and extinction. *Psychological review*, 117(1):197, 2010.

S. J. Gershman, K. A. Norman, and Y. Niv. Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, 5:43–50, 2015. ISSN 2352-1546. doi: https:

//doi.org/10.1016/j.cobeha.2015.07.007. URL https://www.sciencedirect.com/science/article/pii/S2352154615001059.

M. Geva-Sagiv, E. A. Mankin, D. Eliashiv, S. Epstein, N. Cherry, G. Kalender, N. Tchemodanov, Y. Nir, and I. Fried. Augmenting hippocampal–prefrontal neuronal synchrony during sleep enhances memory consolidation in humans. *Nature neuroscience*, 26 (6):1100–1110, 2023.

A. K. Gillespie, D. A. A. Maya, E. L. Denovellis, D. F. Liu, D. B. Kastner, M. E. Coulter, D. K. Roumis, U. T. Eden, and L. M. Frank. Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, 109(19):3149–3163, 2021.

G. Girardeau, K. Benchenane, S. I. Wiener, G. Buzsáki, and M. B. Zugaro. Selective suppression of hippocampal ripples impairs spatial memory. *Nature neuroscience*, 12(10):1222–1223, 2009.

O. C. González, Y. Sokolov, G. P. Krishnan, J. E. Delanois, and M. Bazhenov. Can sleep protect memories from catastrophic forgetting? *Elife*, 9:e51005, 2020.

K. Gorgolewski, C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5:13, 2011. doi: 10.3389/fninf.2011.00013.

K. J. Gorgolewski, O. Esteban, C. J. Markiewicz, E. Ziegler, D. G. Ellis, M. P. Notter, D. Jarecka, H. Johnson, C. Burns, A. Manhães-Savio, C. Hamalainen, B. Yvernault, T. Salo, K. Jordan, M. Goncalves, M. Waskom, D. Clark, J. Wong, F. Loney, M. Modat, B. E. Dewey, C. Madison, M. Visconti di Oleggio Castello, M. G. Clark, M. Dayan, D. Clark, A. Keshavan, B. Pinsard, A. Gramfort, S. Berleant, D. M. Nielson, S. Bougacha, G. Varoquaux, B. Cipollini, R. Markello, A. Rokem, B. Moloney, Y. O. Halchenko, D. Wassermann, M. Hanke, C. Horea, J. Kaczmarzyk, G. de Hollander, E. DuPre, A. Gillman, D. Mordom, C. Buchanan, R. Tungaraza, W. M. Pauli, S. Iqbal, S. Sikka, M. Mancini, Y. Schwartz, I. B. Malone, M. Dubois, C. Frohlich, D. Welch, J. Forbes, J. Kent, A. Watanabe, C. Cumba, J. M. Huntenburg, E. Kastman, B. N. Nichols, A. Eshaghi, D. Ginsburg, A. Schaefer, B. Acland, S. Giavasis, J. Kleesiek, D. Erickson, R. Küttner, C. Haselgrove, C. Correa, A. Ghayoor, F. Liem, J. Millman, D. Haehn, J. Lai, D. Zhou, R. Blair, T. Glatard, M. Renfro, S. Liu, A. E. Kahn, F. Pérez-García, W. Triplett, L. Lampe, J. Stadler, X.-Z. Kong, M. Hallquist, A. Chetverikov, J. Salvatore, A. Park, R. Poldrack, R. C. Craddock, S. Inati, O. Hinds, G. Cooper, L. N. Perkins, A. Marina, A. Mattfeld, M. Noel, L. Snoek, K. Matsubara, B. Cheung, S. Rothmei, S. Urchs, J. Durnez, F. Mertz, D. Geisler, A. Floren, S. Gerhard, P. Sharp, M. Molina-Romero, A. Weinstein, W. Broderick, V. Saase, S. K. Andberg, R. Harms, K. Schlamp, J. Arias, D. Papadopoulos Orfanos, C. Tarbert, A. Tambini, A. De La Vega,

T. Nickson, M. Brett, M. Falkiewicz, K. Podranski, J. Linkersdörfer, G. Flandin, E. Ort, D. Shachnev, D. McNamee, A. Davison, J. Varada, I. Schwabacher, J. Pellman, M. Perez-Guevara, R. Khanuja, N. Pannetier, C. McDermottroe, and S. Ghosh. Nipype. *Software*, 2018. doi: 10.5281/zenodo.596855.

C. L. Grady and D. D. Garrett. Understanding variability in the BOLD signal and why it matters for aging. *Brain Imaging and Behavior*, 8(2):274–283, June 2014. ISSN 1931-7557, 1931-7565. doi: 10.1007/s11682-013-9253-0. URL http://link.springer.com/10.1007/s11682-013-9253-0.

A. Graves, G. Wayne, M. Reynolds, T. Harley, I. Danihelka, A. Grabska-Barwińska, S. G. Colmenarejo, E. Grefenstette, T. Ramalho, J. Agapiou, et al. Hybrid computing using a neural network with dynamic external memory. *Nature*, 538(7626):471–476, 2016.

D. N. Greve and B. Fischl. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72, 2009. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2009.06.060.

T. L. Griffiths and Z. Ghahramani. The indian buffet process: An introduction and review. *Journal of Machine Learning Research*, 12(4), 2011.

K. Grill-Spector and R. Malach. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta psychologica*, 107(1-3):293–321, 2001. Publisher: Elsevier.

A. D. Grosmark and G. Buzsáki. Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science*, 351(6280):1440–1443, 2016.

B. Guerra-Carrillo, A. P. Mackey, and S. A. Bunge. Resting-State fMRI: A Window into Human Brain Plasticity. *The Neuroscientist*, 20(5):522–533, Oct. 2014. ISSN 1073-8584, 1089-4098. doi: 10.1177/1073858414524442. URL http://journals.sagepub.com/doi/10.1177/1073858414524442.

A. S. Gupta, M. A. Van Der Meer, D. S. Touretzky, and A. D. Redish. Hippocampal replay is not a simple function of experience. *Neuron*, 65(5):695–705, 2010.

H. Gärtner, M. Minnerop, P. Pieperhoff, A. Schleicher, K. Zilles, E. Altenmüller, and K. Amunts. Brain morphometry shows effects of long-term musical practice in middle-aged keyboard players. *Frontiers in Psychology*, 4, Sept. 2013. ISSN 1664-1078. doi: 10.3389/fpsyg.2013.00636. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2013.00636/full. Publisher: Frontiers.

T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005.

R. H. Hamilton, E. G. Chrysikou, and B. Coslett. Mechanisms of aphasia recovery after stroke and the role of noninvasive brain stimulation. *Brain and Language*, 118(1):40–50, July 2011. ISSN 0093-934X. doi: 10.1016/j.bandl.2011.02.005. URL https://www.sciencedirect.com/science/article/pii/S0093934X1100037X.

T. A. Hare, C. F. Camerer, and A. Rangel. Self-control in decision-making involves modulation of the vmpfc valuation system. *Science*, 324(5927):646–648, 2009.

E. Hariton and J. J. Locascio. Randomised controlled trials—the gold standard for effectiveness research. *BJOG : an international journal of obstetrics and gynaecology*, 125(13):1716, June 2018. doi: 10.1111/1471-0528.15199. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC6235704/.

G. Hartwigsen. Flexible Redistribution in Cognitive Networks. *Trends in Cognitive Sciences*, 22(8):687–698, Aug. 2018. ISSN 1879-307X. doi: 10.1016/j.tics.2018.05.008.

G. Hartwigsen and L. J. Volz. Probing rapid network reorganization of motor and language functions via neuromodulation and neuroimaging. *NeuroImage*, 224:117449, Jan. 2021. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2020.117449.

X. He, P. K. Büchel, S. Faghel-Soubeyrand, J. Klingspohr, M. S. Kehl, and B. P. Staresina. Sleep strengthens successor representations of learned sequences. *bioRxiv*, pages 2025–06, 2025.

D. O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. Wiley and Sons, New York, 1949. ISBN 978-0-471-36727-7.

C. Hertzog and J. R. Nesselroade. Assessing Psychological Change in Adulthood: An Overview of Methodological Issues. *Psychology and Aging*, 18(4):639–657, Dec. 2003. ISSN 1939-1498, 0882-7974. doi: 10.1037/0882-7974.18.4.639. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/0882-7974.18.4.639.

E. M. Hillman. Coupling Mechanism and Significance of the BOLD Signal: A Status Report. *Annual review of neuroscience*, 37:161–181, July 2014. ISSN 0147-006X. doi: 10.1146/annurev-neuro-071013-014111. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4147398/.

S. J. Holdsworth, R. O'Halloran, and K. Setsompop. The quest for high spatial resolution diffusion-weighted imaging of the human brain in vivo. *NMR in Biomedicine*, 32(4):e4056, 2019. ISSN 1099-1492. doi: 10.1002/nbm.4056. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4056. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.4056.

Y. Isomura, A. Sirota, S. Özen, S. Montgomery, K. Mizuseki, D. A. Henze, and G. Buzsáki. Integration and segregation of activity in entorhinal-hippocampal subregions by neocortical slow oscillations. *Neuron*, 52(5):871–882, 2006.

S. P. Jadhav, G. Rothschild, D. K. Roumis, and L. M. Frank. Coordinated excitation and inhibition of prefrontal ensembles during awake hippocampal sharp-wave ripple events. *Neuron*, 90(1):113–127, 2016.

J. G. Jenkins and K. M. Dallenbach. Obliviscence during sleep and waking. *The American Journal of Psychology*, 35(4):605–612, 1924.

M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2): 825–841, 2002. ISSN 1053-8119. doi: 10.1006/nimg.2002.1132. URL http://www.sciencedirect.com/science/article/pii/S1053811902911328.

D. Ji and M. A. Wilson. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature neuroscience*, 10(1):100–107, 2007.

T. Jones and E. A. Rabiner. The Development, Past Achievements, and Future Directions of Brain PET. *Journal of Cerebral Blood Flow & Metabolism*, 32(7):1426–1454, July 2012. ISSN 0271-678X. doi: 10.1038/jcbfm.2012.20. URL https://doi.org/10.1038/jcbfm.2012.20. Publisher: SAGE Publications Ltd STM.

A. E. Kahn, D. S. Bassett, and N. D. Daw. Trial-by-trial learning of successor representations in human behavior. *PLOS Computational Biology*, 21(11):e1013696, 2025.

R. Kaplan, N. W. Schuck, and C. F. Doeller. The role of mental maps in decision-making. *Trends in Neurosciences*, 40(5):256–259, 2017.

L. A. Karaba, H. L. Robinson, R. E. Harvey, W. Chen, A. Fernandez-Ruiz, and A. Oliva. A hippocampal circuit mechanism to balance memory reactivation during sleep. *Science*, 385 (6710):738–743, 2024.

M. P. Karlsson and L. M. Frank. Awake replay of remote experiences in the hippocampus. *Nature neuroscience*, 12(7):913–918, 2009.

J. Keck, C. Barry, C. F. Doeller, and J. Jost. Impact of symmetry in local learning rules on predictive neural representations and generalization in spatial navigation. *PLOS Computational Biology*, 21(6):e1013056, 2025.

G. Kempermann. What Is Adult Hippocampal Neurogenesis Good for? *Frontiers in Neuroscience*, 16:852680, Apr. 2022. ISSN 1662-453X. doi: 10.3389/fnins.2022.852680. URL https://www.frontiersin.org/articles/10.3389/fnins.2022.852680/full.

M. Keramati, A. Dezfouli, and P. Piray. Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS computational biology*, 7(5):e1002055, 2011.

S. Kern, J. Nagel, L. Wittkuhn, S. Gais, R. Dolan, and G. Feld. Challenges in replay detection by tdlm in post-encoding resting state. *eLife*, 2026. doi: 10.7554/eLife.108023.2. URL http://dx.doi.org/10.7554/eLife.108023.2.

J. J. Kim and M. S. Fanselow. Modality-specific retrograde amnesia of fear. *Science*, 256 (5057):675–677, 1992.

J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13):3521–3526, 2017.

A. Klein, S. S. Ghosh, F. S. Bao, J. Giard, Y. Häme, E. Stavsky, N. Lee, B. Rossa, M. Reuter, E. C. Neto, and A. Keshavan. Mindboggling morphometry of human brains. *PLOS Computational Biology*, 13(2):e1005350, 2017. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1005350. URL http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005350.

J. G. Klinzing, N. Niethard, and J. Born. Mechanisms of systems memory consolidation during sleep. *Nature neuroscience*, 22(10):1598–1610, 2019.

S. Y. Ko, Y. Rong, A. I. Ramsaran, X. Chen, A. J. Rashid, A. J. Mocle, J. Dhaliwal, A. Awasthi, A. Guskjolen, S. A. Josselyn, et al. Systems consolidation reorganizes hippocampal engram circuitry. *Nature*, 643(8072):735–743, 2025.

R. S. Koolschijn, U. E. Emir, A. C. Pantelides, H. Nili, T. E. Behrens, and H. C. Barron. The Hippocampus and Neocortical Inhibitory Engrams Protect against Memory Interference. *Neuron*, 101(3):528–541.e6, Feb. 2019. ISSN 08966273. doi: 10.1016/j.neuron.2018.11.042. URL https://linkinghub.elsevier.com/retrieve/pii/S0896627318310511.

B. J. Kraus, R. J. Robinson, J. A. White, H. Eichenbaum, and M. E. Hasselmo. Hippocampal "time cells": time versus path integration. *Neuron*, 78(6):1090–1101, 2013.

N. Kriegeskorte, M. Mur, and P. A. Bandettini. Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2:249, 2008. Publisher: Frontiers.

E. Kropff, J. E. Carmichael, M.-B. Moser, and E. I. Moser. Speed cells in the medial entorhinal cortex. *Nature*, 523(7561):419–424, 2015.

D. Kumaran, D. Hassabis, and J. L. McClelland. What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends in cognitive sciences*, 20(7): 512–534, 2016.

F. Kurth, E. Luders, and C. Gaser. Voxel-Based Morphometry. In *Brain Mapping*, pages 345–349. Elsevier, 2015. ISBN 978-0-12-397316-0. doi: 10.1016/B978-0-12-397025-1.00304-3. URL https://linkinghub.elsevier.com/retrieve/pii/B9780123970251003043.

Z. Kurth-Nelson, M. Economides, R. J. Dolan, and P. Dayan. Fast sequences of non-spatial state representations in humans. *Neuron*, 91(1):194–204, 2016.

Z. Kurth-Nelson, T. Behrens, G. Wayne, K. Miller, L. Luettgau, R. Dolan, Y. Liu, and P. Schwartenbeck. Replay and compositional computation. *Neuron*, 111(4):454–469, Feb. 2023. ISSN 08966273. doi: 10.1016/j.neuron.2022.12.028. URL https://linkinghub.elsevier.com/retrieve/pii/S0896627322011254.

C. Lanczos. Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1):76–85, 1964. ISSN 0887-459X. doi: 10.1137/0701007. URL http://epubs.siam.org/doi/10.1137/0701007.

A. K. Lee and M. A. Wilson. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*, 36(6):1183–1194, 2002.

S. W. Lee, S. Shimojo, and J. P. O'doherty. Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, 81(3):687–699, 2014.

J.-P. Lefaucheur, A. Antal, R. Ahdab, D. Ciampi de Andrade, F. Fregni, E. M. Khedr, M. Nitsche, and W. Paulus. The use of repetitive transcranial magnetic stimulation (rTMS) and transcranial direct current stimulation (tDCS) to relieve pain. *Brain Stimulation*, 1(4): 337–344, Oct. 2008. ISSN 1935-861X. doi: 10.1016/j.brs.2008.07.003. URL https://www.sciencedirect.com/science/article/pii/S1935861X08003215.

L. Lehnert, M. L. Littman, and M. J. Frank. Reward-predictive representations generalize across tasks in reinforcement learning. *PLOS Computational Biology*, 16(10):e1008317, Oct. 2020. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1008317. URL https://dx.plos.org/10.1371/journal.pcbi.1008317.

L. Li, T. J. Walsh, and M. L. Littman. Towards a unified theory of state abstraction for mdps. *AI&M*, 1(2):3, 2006.

U. Lindenberger, T. Von Oertzen, P. Ghisletta, and C. Hertzog. Cross-sectional age variance extraction: What's change got to do with it? *Psychology and Aging*, 26(1):34–47, 2011.

ISSN 1939-1498, 0882-7974. doi: 10.1037/a0020525. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0020525.

Y. Liu, R. J. Dolan, Z. Kurth-Nelson, and T. E. Behrens. Human replay spontaneously reorganizes experience. *Cell*, 178(3):640–652, 2019.

Y. Liu, R. J. Dolan, C. Higgins, H. Penagos, M. W. Woolrich, H. F. Ólafsdóttir, C. Barry, Z. Kurth-Nelson, and T. E. Behrens. Temporally delayed linear modelling (tdlm) measures replay in both animals and humans. *Elife*, 10:e66917, 2021a.

Y. Liu, M. G. Mattar, T. E. Behrens, N. D. Daw, and R. J. Dolan. Experience replay is associated with efficient nonlocal learning. *Science*, 372(6544):eabf1357, 2021b.

Y. Liu, M. M. Nour, N. W. Schuck, T. E. Behrens, and R. J. Dolan. Decoding cognition from spontaneous neural activity. *Nature Reviews Neuroscience*, 23(4):204–214, 2022.

K. Lloyd and D. S. Leslie. Context-dependent decision-making: a simple bayesian model. *Journal of The Royal Society Interface*, 10(82), 2013.

A. T. Löwe, M. Petzka, M. M. Tzegka, and N. W. Schuck. N2 sleep promotes the occurrence of 'aha'moments in a perceptual insight task. *PLoS Biology*, 23(6):e3003185, 2025.

N. D. Lutz, M. Harkotte, and J. Born. Sleep's contribution to memory formation. *Physiological Reviews*, 106(1):363–483, 2026.

V. López-Alonso, B. Cheeran, D. Río-Rodríguez, and M. Fernández-del Olmo. Inter-individual Variability in Response to Non-invasive Brain Stimulation Paradigms. *Brain Stimulation*, 7(3):372–380, May 2014. ISSN 1935-861X. doi: 10.1016/j.brs.2014.02.004. URL https://www.sciencedirect.com/science/article/pii/S1935861X14000989.

M. Lövdén, N. C. Bodammer, S. Kühn, J. Kaufmann, H. Schütze, C. Tempelmann, H.-J. Heinze, E. Düzel, F. Schmiedek, and U. Lindenberger. Experience-dependent plasticity of white-matter microstructure extends into old age. *Neuropsychologia*, 48(13):3878–3883, Nov. 2010a. ISSN 1873-3514. doi: 10.1016/j.neuropsychologia.2010.08.026.

M. Lövdén, L. Bäckman, U. Lindenberger, S. Schaefer, and F. Schmiedek. A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, 136(4):659–676, 2010b. ISSN 1939-1455, 0033-2909. doi: 10.1037/a0020080. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0020080.

C. J. MacDonald, K. Q. Lepage, U. T. Eden, and H. Eichenbaum. Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, 71(4):737–749, 2011.

E. A. Maguire, D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. J. Frackowiak, and C. D. Frith. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97(8):4398–4403, Apr. 2000. doi: 10.1073/pnas.070039597. URL https://www.pnas.org/doi/10.1073/pnas.070039597. Publisher: Proceedings of the National Academy of Sciences.

E. A. Maguire, K. Woollett, and H. J. Spiers. London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus*, 16(12):1091–1101, 2006. Publisher: Wiley Online Library.

R. C. Malenka and M. F. Bear. LTP and LTD: An Embarrassment of Riches. *Neuron*, 44(1): 5–21, Sept. 2004. ISSN 0896-6273. doi: 10.1016/j.neuron.2004.09.012. URL https://www.cell.com/neuron/abstract/S0896-6273(04)00608-7. Publisher: Elsevier.

D. Marr. Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 262(841):23–81, 1971.

L. Marshall, M. Mölle, H. R. Siebner, and J. Born. Bifrontal transcranial direct current stimulation slows reaction time in a working memory task. *BMC Neuroscience*, 6(1):23, Apr. 2005. ISSN 1471-2202. doi: 10.1186/1471-2202-6-23. URL https://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-6-23.

M. G. Mattar and N. D. Daw. Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, 21(11):1609–1617, Nov. 2018. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-018-0232-z. URL https://www.nature.com/articles/s41593-018-0232-z.

W. Mau, M. E. Hasselmo, and D. J. Cai. The brain in motion: How ensemble fluidity drives memory-updating and flexibility. *Elife*, 9:e63550, 2020.

T. Maviel, T. P. Durkin, F. Menzaghi, and B. Bontempi. Sites of neocortical reorganization critical for remote spatial memory. *Science*, 305(5680):96–99, 2004.

R. A. McCallum. Instance-based utile distinctions for reinforcement learning with hidden state. In *Machine Learning Proceedings 1995*, pages 387–395. Elsevier, 1995.

J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3):419, 1995.

J. L. McClelland, B. L. McNaughton, and A. K. Lampinen. Integration of new information in memory: new insights from a complementary learning systems perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1799), 2020.

M. McCloskey and N. J. Cohen. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, pages 109–165. Elsevier, 1989.

J. A. Miller and C. Constantinidis. Timescales of learning in prefrontal cortex. *Nature Reviews Neuroscience*, June 2024. ISSN 1471-003X, 1471-0048. doi: 10.1038/s41583-024-00836-8. URL https://www.nature.com/articles/s41583-024-00836-8.

D.-M. Mirea, Y. S. Shin, S. DuBrow, and Y. Niv. The ubiquity of time in latent-cause inference. *Journal of cognitive neuroscience*, 36(11):2442–2454, 2024.

V. Mnih, N. Heess, A. Graves, and K. Kavukcuoglu. Recurrent models of visual attention. *Advances in neural information processing systems*, 27, 2014.

I. Momennejad, E. M. Russek, J. H. Cheong, M. M. Botvinick, N. D. Daw, and S. J. Gershman. The successor representation in human reinforcement learning. *Nature Human Behaviour*, 1(9):680–692, Aug. 2017. ISSN 2397-3374. doi: 10.1038/s41562-017-0180-8. URL https://www.nature.com/articles/s41562-017-0180-8.

N. Moneta, S. Grossman, and N. W. Schuck. Representational spaces in orbitofrontal and ventromedial prefrontal cortex: task states, values, and beyond. *Trends in Neurosciences*, 47 (12):1055–1069, 2024.

P. R. Montague, P. Dayan, and T. J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of neuroscience*, 16(5):1936–1947, 1996.

K. Monte-Silva, M.-F. Kuo, S. Hessenthaler, S. Fresnoza, D. Liebetanz, W. Paulus, and M. A. Nitsche. Induction of Late LTP-Like Plasticity in the Human Motor Cortex by Repeated Non-Invasive Brain Stimulation. *Brain Stimulation*, 6(3):424–432, May 2013. ISSN 1935-861X. doi: 10.1016/j.brs.2012.04.011. URL https://www.sciencedirect.com/science/article/pii/S1935861X12000794.

T. D. Moody, S. Y. Bookheimer, Z. Vanek, and B. J. Knowlton. An Implicit Learning Task Activates Medial Temporal Lobe in Patients With Parkinson's Disease. *Behavioral Neuroscience*, 118(2):438–442, 2004. ISSN 1939-0084, 0735-7044. doi: 10.1037/0735-7044.118.2.438. URL https://doi.apa.org/doi/10.1037/0735-7044.118.2.438. Publisher: American Psychological Association (APA).

J. H. Morrison and M. G. Baxter. The ageing cortical synapse: hallmarks and implications for cognitive decline. *Nature Reviews Neuroscience*, 13(4):240–250, Apr. 2012. ISSN 1471-003X, 1471-0048. doi: 10.1038/nrn3200. URL https://www.nature.com/articles/nrn3200.

L. Mosconi. Brain glucose metabolism in the early and specific diagnosis of Alzheimer?s disease: FDG-PET studies in MCI and AD. *European Journal of Nuclear Medicine and Molecular Imaging*, 32(4):486–510, Apr. 2005. ISSN 1619-7070, 1619-7089. doi: 10.1007/s00259-005-1762-7. URL http://link.springer.com/10.1007/s00259-005-1762-7.

M. Moscovitch, R. Cabeza, G. Winocur, and L. Nadel. Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annual review of psychology*, 67(1):105–134, 2016.

E. I. Moser, E. Kropff, and M.-B. Moser. Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, 31(1):69–89, 2008.

E. I. Moser, M.-B. Moser, and B. L. McNaughton. Spatial representation in the hippocampal formation: a history. *Nature neuroscience*, 20(11):1448–1464, 2017.

Z. Nádasdy, H. Hirase, A. Czurkó, J. Csicsvari, and G. Buzsáki. Replay and time compression of recurring spike sequences in the hippocampus. *Journal of neuroscience*, 19(21):9497–9507, 1999.

L. Nadel and M. Moscovitch. Memory consolidation, retrograde amnesia and the hippocampal complex. *Current opinion in neurobiology*, 7(2):217–227, 1997.

M. R. Nassar, R. C. Wilson, B. Heasly, and J. I. Gold. An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, 30(37):12366–12378, 2010.

D. J. Newbold, T. O. Laumann, C. R. Hoyt, J. M. Hampton, D. F. Montez, R. V. Raut, M. Ortega, A. Mitra, A. N. Nielsen, D. B. Miller, B. Adeyemo, A. L. Nguyen, K. M. Scheidter, A. B. Tanenbaum, A. N. Van, S. Marek, B. L. Schlaggar, A. R. Carter, D. J. Greene, E. M. Gordon, M. E. Raichle, S. E. Petersen, A. Z. Snyder, and N. U. Dosenbach. Plasticity and Spontaneous Activity Pulses in Disused Human Brain Circuits. *Neuron*, 107(3):580–589.e6, Aug. 2020. ISSN 08966273. doi: 10.1016/j.neuron.2020.05.007. URL https://linkinghub.elsevier.com/retrieve/pii/S0896627320303536.

H. Nili, C. Wingfield, A. Walther, L. Su, W. Marslen-Wilson, and N. Kriegeskorte. A toolbox for representational similarity analysis. *PLoS computational biology*, 10(4):e1003553, 2014.

Y. Niv. Learning task-state representations. *Nature neuroscience*, 22(10):1544–1553, 2019.

Y. Niv, R. Daniel, A. Geana, S. J. Gershman, Y. C. Leong, A. Radulescu, and R. C. Wilson. Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, 35(21):8145–8157, 2015.

K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby. Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in cognitive sciences*, 10(9):424–430, 2006.

M. M. Nour, Y. Liu, A. Arumuham, Z. Kurth-Nelson, and R. J. Dolan. Impaired neural replay of inferred relationships in schizophrenia. *Cell*, 184(16):4315–4328, 2021.

J. O'Doherty, P. Dayan, J. Schultz, R. Deichmann, K. Friston, and R. J. Dolan. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *science*, 304(5669):452–454, 2004.

S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, Dec. 1990. doi: 10.1073/pnas.87.24.9868. URL https://www.pnas.org/doi/abs/10.1073/pnas.87.24.9868. Publisher: Proceedings of the National Academy of Sciences.

J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 1971.

J. O'Keefe and L. Nadel. *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press, 1978. ISBN 978-0-19-857206-0. URL https://repository.arizona.edu/handle/10150/620894. Accepted: 2016-10-08T00:43:54Z.

H. F. Ólafsdóttir, F. Carpenter, and C. Barry. Coordinated grid and place cell replay during rest. *Nature neuroscience*, 19(6):792–794, 2016.

H. F. Ólafsdóttir, D. Bush, and C. Barry. The role of hippocampal replay in memory and planning. *Current Biology*, 28(1):R37–R50, 2018.

J. Ou, Y. Qu, Y. Xu, Z. Xiao, T. Behrens, and Y. Liu. Replay builds an efficient cognitive map offline to avoid computation online. *bioRxiv*, pages 2025–01, 2025.

J. Paillard. Reflections on the usage of the concept of plasticity in neurobiology. *Journal de Psychologie Normale et Pathologique*, 73(1):33–47, 1976. ISSN 0021-7956. Place: France Publisher: Presses Universitaires de France.

A. Pascual-Leone, J. R. Gates, and A. Dhuna. Induction of speech arrest and counting errors with rapid-rate transcranial magnetic stimulation. *Neurology*, 41(5):697–702, May 1991. doi: 10.1212/WNL.41.5.697. URL https://www.neurology.org/doi/10.1212/WNL.41.5.697. Publisher: Wolters Kluwer.

R. Patriat, R. C. Reynolds, and R. M. Birn. An improved model of motion-related signal changes in fMRI. *NeuroImage*, 144, Part A:74–82, Jan. 2017. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2016.08.051. URL http://www.sciencedirect.com/science/article/pii/S1053811916304360.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

W. W. Pettine, D. V. Raman, A. D. Redish, and J. D. Murray. Human latent-state generalization through prototype learning with discriminative attention. preprint, PsyArXiv, Dec. 2021. URL https://osf.io/ku4fr.

B. E. Pfeiffer and D. J. Foster. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79, 2013.

P. Piray and N. D. Daw. Linear reinforcement learning in planning, grid fields, and cognitive control. *Nature communications*, 12(1):4942, 2021.

S. Pisupati, I. M. Berwian, J. C. Chiu, Y. Ren, and Y. Niv. Human inductive biases for aversive continual learning-a hierarchical bayesian nonparametric model. In *CoLLAs*, pages 337–346, 2023.

S. Pisupati, A. J. Langdon, A. B. Konova, and Y. Niv. The utility of a latent-cause framework for understanding addiction phenomena. *Addiction neuroscience*, 10:100143, 2024.

R. A. Poldrack. Imaging Brain Plasticity: Conceptual and Methodological Issues— A Theoretical Review. *NeuroImage*, 12(1):1–13, July 2000. ISSN 10538119. doi: 10.1006/nimg.2000.0596. URL https://linkinghub.elsevier.com/retrieve/pii/S1053811900905962.

R. A. Poldrack, J. Clark, E. J. Paré-Blagoev, D. Shohamy, J. Creso Moyano, C. Myers, and M. A. Gluck. Interactive memory systems in the human brain. *Nature*, 414(6863):546–550, Nov. 2001. ISSN 0028-0836, 1476-4687. doi: 10.1038/35107080. URL https://www.nature.com/articles/35107080. Publisher: Springer Science and Business Media LLC.

J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fmri. *NeuroImage*, 84(Supplement C):320–341, 2014. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2013.08.048. URL http://www.sciencedirect.com/science/article/pii/S1053811913009117.

A. R. Powers, M. A. Hevey, and M. T. Wallace. Neural correlates of multisensory perceptual learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(18):6263–6274, May 2012. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.6138-11.2012.

P. P. Qin, M. Jin, A. W. Xia, A. S. Li, T. T. Lin, Y. Liu, R. L. Kan, B. B. Zhang, and G. S. Kranz. The effectiveness and safety of low-intensity transcranial ultrasound stimulation: A systematic review of human and animal studies. *Neuroscience & Biobehavioral Reviews*, 156:105501, Jan. 2024. ISSN 0149-7634. doi: 10.1016/j.neubiorev.2023.105501. URL https://www.sciencedirect.com/science/article/pii/S0149763423004700.

A. Radulescu, Y. S. Shin, and Y. Niv. Human representation learning. *Annual Review of Neuroscience*, 44(1):253–273, 2021.

S. Ramón y Cajal. The croonian lecture.—la fine structure des centres nerveux. *Proceedings of the Royal Society of London*, 55(331-335):444–468, 1894.

B. Rasch and J. Born. About sleep's role in memory. *Physiological reviews*, 2013.

B. Rasch, C. Buchel, S. Gais, and J. Born. Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, 315(5817):1426–1429, 2007.

C. Rasmussen. The infinite gaussian mixture model. *Advances in neural information processing systems*, 12, 1999.

N. Raz, U. Lindenberger, K. M. Rodrigue, K. M. Kennedy, D. Head, A. Williamson, C. Dahle, D. Gerstorf, and J. D. Acker. Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*, 15 (11):1676–1689, Nov. 2005. ISSN 1460-2199, 1047-3211. doi: 10.1093/cercor/bhi044. URL http://academic.oup.com/cercor/article/15/11/1676/296890/Regional-Brain-Changes-in-Aging-Healthy-Adults.

F. M. Renz, S. Grossman, N. D. Daw, P. Dayan, C. F. Doeller, and N. W. Schuck. Neural replay is connected to latent cause inference and supports fast generalization. *bioRxiv*, 2026. doi: 10.64898/2025.12.12.693963. URL https://www.biorxiv.org/content/early/2026/01/16/2025.12.12.693963.

R. A. Rescorla. A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. *Classical conditioning, Current research and theory*, 2: 64–69, 1972.

M. Reuter, H. D. Rosas, and B. Fischl. Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, 53(4):1181–1196, 2010. doi: 10.1016/j.neuroimage.2010.07.020.

F. R. Richter, R. A. Cooper, P. M. Bays, and J. S. Simons. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *elife*, 5:e18260, 2016.

J. Robin and M. Moscovitch. Details, gist and schema: hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. *Current opinion in behavioral sciences*, 17:114–123, 2017.

J. A. Rozo, I. Martínez-Gallego, and A. Rodríguez-Moreno. Cajal, the neuronal theory and the idea of brain plasticity. *Frontiers in neuroanatomy*, 18:1331666, 2024.

M. D. Rugg and D. R. King. Ventral lateral parietal cortex and episodic memory retrieval. *Cortex*, 107:238–250, 2018.

M. E. Rule, T. O'Leary, and C. D. Harvey. Causes and consequences of representational drift. *Current opinion in neurobiology*, 58:141–147, 2019.

E. M. Russek, I. Momennejad, M. M. Botvinick, S. J. Gershman, and N. D. Daw. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS computational biology*, 13(9):e1005768, 2017.

E. M. Russek, I. Momennejad, M. M. Botvinick, S. J. Gershman, and N. D. Daw. Neural evidence for the successor representation in choice evaluation. preprint, Neuroscience, Aug. 2021. URL http://biorxiv.org/lookup/doi/10.1101/2021.08.29.458114.

Y. Sagiv, T. Akam, I. B. Witten, and N. D. Daw. Between planning and map building: Prioritizing replay when future goals are uncertain. *Neuron*, 113(24):4278–4292, 2025.

C. Sampaio-Baptista, Z.-B. Sanders, and H. Johansen-Berg. Structural Plasticity in Adulthood with Motor Learning and Stroke Rehabilitation. *Annual Review of Neuroscience*, 41(Volume 41, 2018):25–40, July 2018. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev-neuro-080317-062015. URL https://www.annualreviews.org/content/journals/10.1146/annurev-neuro-080317-062015. Publisher: Annual Reviews.

A. N. Sanborn, T. L. Griffiths, and D. J. Navarro. Rational approximations to rational models: alternative algorithms for category learning. *Psychological review*, 117(4):1144, 2010.

K. Sato, E. Kirino, and S. Tanaka. A Voxel-Based Morphometry Study of the Brain of University Students Majoring in Music and Nonmusic Disciplines. *Behavioural Neurology*, 2015: 274919, 2015. ISSN 1875-8584. doi: 10.1155/2015/274919.

T. D. Satterthwaite, M. A. Elliott, R. T. Gerraty, K. Ruparel, J. Loughead, M. E. Calkins, S. B. Eickhoff, H. Hakonarson, R. C. Gur, R. E. Gur, and D. H. Wolf. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, 64(1):240–256, 2013.

ISSN 10538119. doi: 10.1016/j.neuroimage.2012.08.052. URL http://linkinghub.elsevier.com/retrieve/pii/S1053811912008609.

C. M. Saucedo Marquez, X. Zhang, S. P. Swinnen, R. Meesen, and N. Wenderoth. Task-Specific Effect of Transcranial Direct Current Stimulation on Motor Learning. *Frontiers in Human Neuroscience*, 7, July 2013. ISSN 1662-5161. doi: 10.3389/fnhum.2013.00333. URL https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2013.00333/full. Publisher: Frontiers.

A. C. Schapiro, T. T. Rogers, N. I. Cordova, N. B. Turk-Browne, and M. M. Botvinick. Neural representations of events arise from temporal community structure. *Nature neuroscience*, 16 (4):486–492, 2013.

A. C. Schapiro, N. B. Turk-Browne, M. M. Botvinick, and K. A. Norman. Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 2017.

F. Scharnowski, R. Veit, R. Zopf, P. Studer, S. Bock, J. Diedrichsen, R. Goebel, K. Mathiak, N. Birbaumer, and N. Weiskopf. Manipulating motor performance and memory through real-time fMRI neurofeedback. *Biological psychology*, 108:85–97, 2015. URL https://www.sciencedirect.com/science/article/pii/S0301051115000678. Publisher: Elsevier.

T. Schaul, J. Quan, I. Antonoglou, and D. Silver. Prioritized experience replay. *arXiv preprint arXiv:1511.05952*, 2015.

E. Schechtman. When memories get complex, sleep comes to their rescue. *Proceedings of the National Academy of Sciences*, 121(12):e2402178121, 2024.

G. Schlaug, L. Jäncke, Y. Huang, J. F. Staiger, and H. Steinmetz. Increased corpus callosum size in musicians. *Neuropsychologia*, 33(8):1047–1055, Aug. 1995. ISSN 00283932. doi: 10.1016/0028-3932(95)00045-5. URL https://linkinghub.elsevier.com/retrieve/pii/0028393295000455.

N. W. Schuck and Y. Niv. Sequential replay of nonspatial task states in the human hippocampus. *Science*, 364(6447):eaaw5181, 2019.

N. W. Schuck, C. F. Doeller, T. A. Polk, U. Lindenberger, and S.-C. Li. Human aging alters the neural computation and representation of space. *NeuroImage*, 117:141–150, Aug. 2015a. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2015.05.031. URL https://www.sciencedirect.com/science/article/pii/S1053811915004115.

N. W. Schuck, R. Gaschler, D. Wenke, J. Heinzle, P. A. Frensch, J.-D. Haynes, and C. Reverberi. Medial Prefrontal Cortex Predicts Internally Driven Strategy Shifts. *Neuron*, 86(1):331–340, Apr. 2015b. ISSN 0896-6273. doi: 10.1016/j.neuron.2015.03.015. URL https://www.cell.com/neuron/abstract/S0896-6273(15)00211-1. Publisher: Elsevier.

N. W. Schuck, M. B. Cai, R. C. Wilson, and Y. Niv. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6):1402–1412, 2016.

W. Schultz. Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, 1998.

W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997.

P. Schwartenbeck, A. Baram, Y. Liu, S. Mark, T. Muller, R. Dolan, M. Botvinick, Z. Kurth-Nelson, and T. Behrens. Generative replay underlies compositional inference in the hippocampal-prefrontal circuit. *Cell*, 186(22):4885–4897, 2023.

W. B. Scoville and B. Milner. Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, 20(1):11, 1957.

M. J. Sekeres, M. Moscovitch, and G. Winocur. Mechanisms of memory consolidation and transformation. In *Cognitive neuroscience of memory consolidation*, pages 17–44. Springer, 2017.

R. N. Shepard. Toward a Universal Law of Generalization for Psychological Science. *Science*, 237(4820):1317–1323, Sept. 1987. doi: 10.1126/science.3629243. URL https://www.science.org/doi/10.1126/science.3629243. Publisher: American Association for the Advancement of Science.

K. Shibata, L.-H. Chang, D. Kim, J. E. N. Sr, Y. Kamitani, T. Watanabe, and Y. Sasaki. Decoding Reveals Plasticity in V3A as a Result of Motion Perceptual Learning. *PLOS ONE*, 7(8):e44003, Aug. 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0044003. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0044003. Publisher: Public Library of Science.

A. P. Shimamura. Episodic retrieval and the cortical binding of relational activity. *Cognitive, Affective, & Behavioral Neuroscience*, 11(3):277–291, 2011.

H. R. Siebner, K. Funke, A. S. Aberra, A. Antal, S. Bestmann, R. Chen, J. Classen, M. Davare, V. Di Lazzaro, P. T. Fox, M. Hallett, A. N. Karabanov, J. Kesselheim, M. M. Beck, G. Koch, D. Liebetanz, S. Meunier, C. Miniussi, W. Paulus, A. V. Peterchev, T. Popa, M. C. Ridding, A. Thielscher, U. Ziemann, J. C. Rothwell, and Y. Ugawa. Transcranial magnetic stimulation

of the brain: What is stimulated? - A consensus and critical position paper. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 140:59–97, Aug. 2022. ISSN 1872-8952. doi: 10.1016/j.clinph.2022.04.022.

E. M. Siefert, S. Uppuluri, J. Mu, M. C. Tandoc, J. W. Antony, and A. C. Schapiro. Memory reactivation during sleep does not act holistically on object memory. *Journal of Neuroscience*, 44(24), 2024.

A. H. Sinclair and M. D. Barense. Prediction error and memory reactivation: How incomplete reminders drive reconsolidation. *Trends in neurosciences*, 42(10):727–739, 2019.

D. Singh, K. A. Norman, and A. C. Schapiro. A model of autonomous interactions between hippocampus and neocortex driving sleep-dependent memory consolidation. *Proceedings of the national academy of sciences*, 119(44):e2123432119, 2022.

A. Sirota, J. Csicsvari, D. Buhl, and G. Buzsáki. Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences*, 100(4):2065–2069, 2003.

R. Sitaram, T. Ros, L. Stoeckel, S. Haller, F. Scharnowski, J. Lewis-Peacock, N. Weiskopf, M. L. Blefari, M. Rana, and E. Oblak. Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, 18(2):86–100, 2017. URL https://www.nature.com/articles/nrn.2016.164. Publisher: Nature Publishing Group UK London.

W. E. Skaggs and B. L. McNaughton. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, 271(5257):1870–1873, 1996.

M. W. Sliwinska, S. Vitello, and J. T. Devlin. Transcranial magnetic stimulation for investigating causal brain-behavioral relationships and their time course. *Journal of Visualized Experiments: JoVE*, (89):51735, July 2014. ISSN 1940-087X. doi: 10.3791/51735.

J. M. Soares, P. Marques, V. Alves, and N. Sousa. A hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neuroscience*, 7, 2013. ISSN 1662-453X. doi: 10.3389/fnins.2013.00031. URL http://journal.frontiersin.org/article/10.3389/fnins.2013.00031/abstract.

S. F. Sorrells, M. F. Paredes, A. Cebrian-Silla, K. Sandoval, D. Qi, K. W. Kelley, D. James, S. Mayer, J. Chang, K. I. Auguste, E. F. Chang, A. J. Gutierrez, A. R. Kriegstein, G. W. Mathern, M. C. Oldham, E. J. Huang, J. M. Garcia-Verdugo, Z. Yang, and A. Alvarez-Buylla. Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature*, 555(7696):377–381, Mar. 2018. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature25975. URL https://www.nature.com/articles/nature25975.

E. Spens and N. Burgess. A generative model of memory construction and consolidation. *Nature human behaviour*, 8(3):526–543, 2024.

K. L. Stachenfeld, M. M. Botvinick, and S. J. Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653, Nov. 2017. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4650. URL https://www.nature.com/articles/nn.4650.

J. Sulzer, S. Haller, F. Scharnowski, N. Weiskopf, N. Birbaumer, M. L. Blefari, A. B. Bruehl, L. G. Cohen, R. C. DeCharms, and R. Gassert. Real-time fMRI neurofeedback: progress and challenges. *Neuroimage*, 76:386–399, 2013. URL https://www.sciencedirect.com/science/article/pii/S1053811913002759. Publisher: Elsevier.

R. S. Sutton. Dyna, an integrated architecture for learning, planning, and reacting. *ACM Sigart Bulletin*, 2(4):160–163, 1991.

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA, 2018. ISBN 0262039249.

R. S. Sutton, D. Precup, and S. Singh. Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, 112(1-2):181–211, 1999.

R. A. Swanson, E. Chinigò, D. Levenstein, M. Vöröslakos, N. Mousavi, X.-J. Wang, J. Basu, and G. Buzsáki. Topography of putative bi-directional interaction between hippocampal sharp-wave ripples and neocortical slow oscillations. *Neuron*, 113(5):754–768, 2025.

A. Takashima, I. L. Nieuwenhuis, O. Jensen, L. M. Talamini, M. Rijpkema, and G. Fernández. Shift from hippocampal to neocortical centered retrieval network with consolidation. *Journal of Neuroscience*, 29(32):10087–10093, 2009.

A. Tambini and L. Davachi. Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends in cognitive sciences*, 23(10):876–890, 2019.

A. A. Taren, P. J. Gianaros, C. M. Greco, E. K. Lindsay, A. Fairgrieve, K. W. Brown, R. K. Rosen, J. L. Ferris, E. Julson, A. L. Marsland, and J. D. Creswell. Mindfulness Meditation Training and Executive Control Network Resting State Functional Connectivity: A Randomized Controlled Trial. *Psychosomatic Medicine*, 79(6): 674, Aug. 2017. ISSN 0033-3174. doi: 10.1097/PSY.0000000000000466. URL https://journals.lww.com/psychosomaticmedicine/abstract/2017/07000/mindfulness_meditation_training_and_executive.10.aspx.

L. Tarhan and T. Konkle. Reliability-based voxel selection. *NeuroImage*, 207:116350, 2020.

J. S. Taube, R. U. Muller, and J. B. Ranck. Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *Journal of Neuroscience*, 10(2):420–435, 1990.

R. M. Tavares, A. Mendelsohn, Y. Grossman, C. H. Williams, M. Shapiro, Y. Trope, and D. Schiller. A map for social navigation in the human brain. *Neuron*, 87(1):231–243, 2015.

D. Tingley and A. Peyrache. On the methods for reactivation and replay analysis. *Philosophical Transactions of the Royal Society B*, 375(1799):20190231, 2020.

E. C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948. ISSN 1939-1471. doi: 10.1037/h0061626. Place: US Publisher: American Psychological Association.

A. Tompary and L. Davachi. Consolidation promotes the emergence of representational overlap in the hippocampus and medial prefrontal cortex. *Neuron*, 96(1):228–241, 2017.

A. Tompary and L. Davachi. Integration of overlapping sequences emerges with consolidation through medial prefrontal cortex neural ensembles and hippocampal–cortical connectivity. *Elife*, 13:e84359, 2024.

A. Tompary and S. L. Thompson-Schill. Semantic influences on episodic memory distortions. *Journal of Experimental Psychology: General*, 150(9):1800, 2021.

D. Tse, R. F. Langston, M. Kakeyama, I. Bethus, P. A. Spooner, E. R. Wood, M. P. Witter, and R. G. Morris. Schemas and memory consolidation. *Science*, 316(5821):76–82, 2007.

N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6): 1310–1320, 2010. ISSN 0278-0062. doi: 10.1109/TMI.2010.2046908.

K. Uğurbil. Magnetic Resonance Imaging at Ultrahigh Fields. *IEEE Transactions on Biomedical Engineering*, 61(5):1364–1379, May 2014. ISSN 1558-2531. doi: 10.1109/TBME.2014.2313619. URL https://ieeexplore.ieee.org/abstract/document/6778087?casa_token=7PaIeBC_io0AAAAA:TFPMzxqw8y4Q3cGnYHIC67rbvK5XG2ygQLsnzfG_OwRqnuiZS87HDTdS0fn5KWkGMmXc_DLjzA. Conference Name: IEEE Transactions on Biomedical Engineering.

K. Uğurbil. ULTRAHIGH FIELD and ULTRAHIGH RESOLUTION fMRI. *Current opinion in biomedical engineering*, 18:100288, Apr. 2021. doi: 10.1016/j.cobme.2021.100288. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC8112570/.

S. Vahdat, M. Darainy, T. E. Milner, and D. J. Ostry. Functionally Specific Changes in Resting-State Sensorimotor Networks after Motor Learning. *Journal of Neuroscience*, 31(47):16907–16915, Nov. 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2737-11.2011. URL https://www.jneurosci.org/content/31/47/16907. Publisher: Society for Neuroscience Section: Articles.

A. R. Vaidya and D. Badre. Abstract task representations for inference and control. *Trends in Cognitive Sciences*, 26(6):484–498, June 2022. ISSN 13646613. doi: 10.1016/j.tics.2022.03.009. URL https://linkinghub.elsevier.com/retrieve/pii/S1364661322000675.

V. V. Valentin, A. Dickinson, and J. P. O'Doherty. Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience*, 27(15):4019–4026, 2007.

M. A. van Der Meer and D. Bendor. Awake replay: off the clock but on the job. *Trends in Neurosciences*, 2025.

A. Vergallito, S. Feroldi, A. Pisoni, and L. J. Romero Lauro. Inter-Individual Variability in tDCS Effects: A Narrative Review on the Contribution of Stable, Variable, and Contextual Factors. *Brain Sciences*, 12(5):522, Apr. 2022. ISSN 2076-3425. doi: 10.3390/brainsci12050522.

L. Verra, F. Renz, and N. W. Schuck. Neuroscience methods for investigating brain plasticity. In A. K. Barbey, editor, *The Oxford Handbook of Cognitive Enhancement and Brain Plasticity*. Oxford University Press, Oxford, 2025. doi: 10.1093/oxfordhb/9780197677131.013.24. URL https://doi.org/10.1093/oxfordhb/9780197677131.013.24. Online edn, Oxford Academic, first published 20 Nov. 2023.

P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

A. N. Voineskos, T. K. Rajji, N. J. Lobaugh, D. Miranda, M. E. Shenton, J. L. Kennedy, B. G. Pollock, and B. H. Mulsant. Age-related decline in white matter tract integrity and cognitive performance: A DTI tractography and structural equation modeling study. *Neurobiology of Aging*, 33(1):21–34, Jan. 2012. ISSN 01974580. doi: 10.1016/j.neurobiolaging.2010.02.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0197458010000874.

N. D. Volkow, Y.-S. Ding, J. S. Fowler, G.-J. Wang, J. Logan, S. J. Gatley, R. Hitzemann, G. Smith, S. D. Fields, and R. Gur. Dopamine transporters decrease with age. *Journal of Nuclear Medicine*, 37(4):554–559, 1996. Publisher: Soc Nuclear Med.

M. W. Voss, R. S. Prakash, K. I. Erickson, C. Basak, L. Chaddock, J. S. Kim, H. Alves, S. Heo, A. Szabo, S. M. White, et al. Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. *Frontiers in aging neuroscience*, 2:1803, 2010.

U. Wagner, S. Gais, H. Haider, R. Verleger, and J. Born. Sleep inspires insight. *Nature*, 427 (6972):352–355, 2004.

A. Walther, H. Nili, N. Ejaz, A. Alink, N. Kriegeskorte, and J. Diedrichsen. Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage*, 137:188–200, 2016.

C. J. Watkins and P. Dayan. Q-learning. *Machine learning*, 8(3):279–292, 1992.

E. Wenger, S. Kühn, J. Verrel, J. Mårtensson, N. C. Bodammer, U. Lindenberger, and M. Lövdén. Repeated Structural Imaging Reveals Nonlinear Progression of Experience-Dependent Volume Changes in Human Motor Cortex. *Cerebral Cortex*, 27(5):2911–2925, May 2017. ISSN 1047-3211. doi: 10.1093/cercor/bhw141. URL https://doi.org/10.1093/cercor/bhw141.

P. Werner, H. Barthel, A. Drzezga, and O. Sabri. Current status and future role of brain PET/MRI in clinical and research settings. *European Journal of Nuclear Medicine and Molecular Imaging*, 42(3):512–526, Mar. 2015. ISSN 1619-7089. doi: 10.1007/s00259-014-2970-9. URL https://doi.org/10.1007/s00259-014-2970-9.

J. C. Whittington, T. H. Muller, S. Mark, G. Chen, C. Barry, N. Burgess, and T. E. Behrens. The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell*, 183(5):1249–1263.e23, Nov. 2020. ISSN 00928674. doi: 10.1016/j.cell.2020.10.024. URL https://linkinghub.elsevier.com/retrieve/pii/S009286742031388X.

A. M. Wikenheiser and G. Schoenbaum. Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nature reviews neuroscience*, 17(8):513–523, 2016.

B. Will, J. Dalrymple-Alford, M. Wolff, and J.-C. Cassel. The concept of brain plasticity—Paillard's systemic analysis and emphasis on structure and function (followed by the translation of a seminal paper by Paillard on plasticity). *Behavioural Brain Research*, 192 (1):2–7, Sept. 2008. ISSN 01664328. doi: 10.1016/j.bbr.2007.11.030. URL https://linkinghub.elsevier.com/retrieve/pii/S0166432807006602.

M. A. Wilson and B. L. McNaughton. Reactivation of hippocampal ensemble memories during sleep. *Science*, 265(5172):676–679, 1994.

R. C. Wilson and Y. Niv. Inferring relevance in a changing world. *Frontiers in human neuroscience*, 5:189, 2012.

R. C. Wilson, Y. K. Takahashi, G. Schoenbaum, and Y. Niv. Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81(2):267–279, 2014.

G. E. Wimmer and D. Shohamy. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*, 338(6104):270–273, 2012.

G. E. Wimmer, N. D. Daw, and D. Shohamy. Generalization of value in reinforcement learning by humans: Generalization of value. *European Journal of Neuroscience*, 35(7):1092–1104, Apr. 2012. ISSN 0953816X. doi: 10.1111/j.1460-9568.2012.08017.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1460-9568.2012.08017.x.

G. E. Wimmer, Y. Liu, N. Vehar, T. E. J. Behrens, and R. J. Dolan. Episodic memory retrieval success is associated with rapid replay of episode content. *Nature Neuroscience*, 23(8): 1025–1033, Aug. 2020. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-020-0649-z. URL https://www.nature.com/articles/s41593-020-0649-z.

G. E. Wimmer, Y. Liu, D. C. McNamee, and R. J. Dolan. Distinct replay signatures for prospective decision-making and memory preservation. *Proceedings of the National Academy of Sciences*, 120(6):e2205211120, 2023.

G. Winocur, M. Moscovitch, and B. Bontempi. Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal–neocortical interactions. *Neuropsychologia*, 48(8):2339–2356, 2010.

L. Wittkuhn and N. W. Schuck. Dynamics of fmri patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nature communications*, 12(1):1795, 2021.

L. Wittkuhn, S. Chien, S. Hall-McMaster, and N. W. Schuck. Replay in minds and machines. *Neuroscience & Biobehavioral Reviews*, 129:367–388, Oct. 2021. ISSN 01497634. doi: 10.1016/j.neubiorev.2021.08.002. URL https://linkinghub.elsevier.com/retrieve/pii/S0149763421003444.

L. Wittkuhn, L. M. Krippner, C. Koch, and N. W. Schuck. Replay in the human visual cortex during brief task pauses is linked to implicit learning of successor representations. *Proceedings of the National Academy of Sciences*, 122(34):e2507516122, 2025.

H. Yuan, K. D. Young, R. Phillips, V. Zotev, M. Misaki, and J. Bodurka. Resting-State Functional Connectivity Modulation and Sustained Changes After Real-Time Functional Magnetic Resonance Imaging Neurofeedback Training in Depression. *Brain Connectivity*, 4(9): 690–701, Nov. 2014. ISSN 2158-0014. doi: 10.1089/brain.2014.0262. URL https://www.liebertpub.com/doi/full/10.1089/brain.2014.0262. Publisher: Mary Ann Liebert, Inc., publishers.

J. M. Zacks and K. M. Swallow. Event segmentation. *Current directions in psychological science*, 16(2):80–84, 2007.

L. B. Zahodne and P. A. Reuter-Lorenz. Compensation and brain aging: A review and analysis of evidence. In *The aging brain: Functional adaptation across adulthood*, pages 185–216. American Psychological Association, Washington, DC, US, 2019. ISBN 978-1-4338-3053-2 978-1-4338-3115-7. doi: 10.1037/0000143-008.

K. Zhang, I. Ginzburg, B. L. McNaughton, and T. J. Sejnowski. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *Journal of neurophysiology*, 79(2):1017–1044, 1998.

Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, 2001. ISSN 0278-0062. doi: 10.1109/42.906424.

J. Zhou, C. Jia, M. Montesinos-Cartagena, M. P. H. Gardner, W. Zong, and G. Schoenbaum. Evolving schema representations in orbitofrontal ensembles during learning. *Nature*, 590(7847):606–611, Feb. 2021. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-020-03061-2. URL https://www.nature.com/articles/s41586-020-03061-2.

W. Zong, J. Zhou, M. P. Gardner, Z. Zhang, K. M. Costa, and G. Schoenbaum. Hippocampal output suppresses orbitofrontal cortex schema cell formation. *Nature Neuroscience*, 28(5):1048–1060, 2025.

R. G. Zsido, A. N. Williams, C. Barth, B. Serio, L. Kurth, T. Mildner, R. Trampel, F. Beyer, A. V. Witte, A. Villringer, and J. Sacher. Ultra-high-field 7T MRI reveals changes in human medial temporal lobe volume in female adults during menstrual cycle. *Nature Mental Health*, 1(10):761–771, Oct. 2023. ISSN 2731-6076. doi: 10.1038/s44220-023-00125-w. URL https://www.nature.com/articles/s44220-023-00125-w. Publisher: Nature Publishing Group.

R. S. Zucker and W. G. Regehr. Short-term synaptic plasticity. *Annual review of physiology*, 64 (1):355–405, 2002. Publisher: Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA.

Acknowledgments

First of all, thank you, dear reader, for taking the time reading this thesis and finding yourself in this section. I hope that you will find something useful in this thesis. If anything sparks your curiosity or if you'd like to discuss any of it further, please don't hesitate to reach out. Looking back, it has always been the interactions with others, whether being advised, bouncing around ideas, or simply talking about science, that have been the most enjoyable moments of this journey. First and foremost, I want to thank my advisors. Nico, I started as a remarkably naive PhD student, unsure of the direction to take, and honestly not really knowing how to do much of anything. Over the course of these years, working closely with you, that has changed, or at least started to. You have taught me a great deal, and I would like to thank you for your patience and guidance. Christian, thank you for the trust and freedom in pursuing my own ideas, and the support in developing my own project, for helping guide it to fruition, and for encouraging me to be ambitious. The experiences I made in Hamburg, Leipzig, and Berlin, but also beyond, in the places I was able to see travelling to Princeton, Oxford, San Francisco, Amsterdam, Barcelona among others, were incredible experiences which I am grateful for. I have been fortunate to have had extraordinary mentors beyond my supervisors. Peter and Nathaniel hosted me in their labs, offering deep insight into the world of computational neuroscience. I learned an enormous amount from both of you, and I very much hope to continue collaborating with and learning from you. I feel like I've only scratched the surface. A PhD is also shaped by the people you share the everyday with. I have been lucky to be embedded in a network of incredibly kind and talented people across labs, cities and programs. In the early days, sitting next to Ondrej, near Christoph, Lennart, Sam, Nir and others who took me under their wing and helped me find my footing. And throughout, from the first attempts at getting pilot experiments running, implementing things in JavaScript fairly miserably, all the way to finishing a project and submitting this thesis, the journey has been a shared one. Shany, I could not have done it without you. Elsa and Marit, thank you for sharing what has been the most ambitious project and data collection I have ever worked on, and for your patience with moody me at 4am. And to my other peers that made the journey what it was: Noa, Lui, Max, Anika, Theo, Alex, Alexa, Judith, Xiangjuan, Neele, Janis and many more. I also want to thank the other colleagues across institutes who made so much of this possible: Gregor, Jöran, Michael and many more. I cannot count the number of times you patiently helped me when I messed things up on the cluster. Thank you for your incredible patience. There are also people who, while not directly involved

in the work itself, have shaped me and accompanied me along the way. Before I even knew where or on what I would do a PhD, people like Frank and Paolo helped shape my thinking and set me on this path. My parents have always made sure that every door I wanted to walk through was open to me, and none of this would have been possible without their support. And friends, Jonas, Tobias, Jeremias, Johannes, Julian, have been indispensable, not just in celebrating milestones but in offering balance, reminding me that it is not all about whether the paper gets published in a certain journal or whether the hypothesis works out. Anne, thank you for being by my side as a partner through all of this, for encouraging me, for helping me navigate the choices and problems that came up along the way. I don't think there is a person who believed more in me than you. We have been an incredible team throughout. To everyone mentioned here, and to those I may have inadvertently left out: thank you for your support and for the time you have given so selflessly. This thesis would not exist without you.

Declaration

Parts of this thesis were written with the assistance of Claude (Anthropic, Claude, Opus 4.6), which was used for improving wording, and refining the writing throughout the drafting process. All scientific content, ideas, and conclusions are my own.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR PSYCHOLOGIE UND
BEWEGUNGSWISSENSCHAFT
Institut für Bewegungswissenschaft
Institut für Psychologie

Eidesstattliche Erklärung nach *(bitte Zutreffendes ankreuzen)*

☐ **§ 7 (4) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☒ **§ 9 (1c und 1d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich an Eides statt,

1. dass die von mir vorgelegte Dissertation nicht Gegenstand eines anderen Prüfungsverfahrens gewesen oder in einem solchen Verfahren als ungenügend beurteilt worden ist.
2. dass ich die von mir vorgelegte Dissertation selbst verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die wörtlich oder inhaltlich übernommenen Stellen habe ich als solche kenntlich gemacht.

Hamburg, 11.4.2026 — Ort, Datum

Renz — Unterschrift

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR PSYCHOLOGIE UND
BEWEGUNGSWISSENSCHAFT
Institut für Bewegungswissenschaft
Institut für Psychologie

Erklärung gemäß *(bitte Zutreffendes ankreuzen)*

☐ **§ 4 (1c) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☒ **§ 5 (4d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich,

Fabian Marvin, Renz (Vorname, Nachname),

dass ich mich an einer anderen Universität oder Fakultät noch keiner Doktorprüfung unterzogen oder mich um Zulassung zu einer Doktorprüfung bemüht habe.

Hamburg, 11.4.2026
Ort, Datum

Renz
Unterschrift